The Formal-Informal Economy Dualism in a Retrospective of Economic Thought since the 1940s

von

Christine Clement

Stuttgart-Hohenheim
ISSN 1618-5358
Unemployment is persistently high in many European countries. It is frequently attributed to ‘functional deficiencies of labor markets’, even though to a large extent caused by demand and supply shifts in goods markets and financial markets. The PhD program Globalization and Employment therefore focuses on the employment effects of structural change and macroeconomic developments and encourages research by granting scholarships. The dissertation projects address topics such as new technologies, processes of internationalization and system transformation, endogenous finance constraints and various strategies of fiscal and monetary policy.

Work and supervision in the PhD program follows the structures of a standard graduate school. Coordinators and editors of the discussion paper series are the following professors:

Prof. Dr. Harald Hagemann  
Department of Economics (520H)  
University of Hohenheim  
D-70593 Stuttgart  
harald.hagemann@uni-hohenheim.de

Prof. Dr. Heinz-Peter Spahn  
Department of Economics (520A)  
University of Hohenheim  
D-70593 Stuttgart  
peter.spahn@uni-hohenheim.de

Prof. Dr. Hans-Michael Trautwein  
Department of Economics and Statistics  
Carl von Ossietzky University Oldenburg  
D-26111 Oldenburg  
michael.trautwein@uni-oldenburg.de

PhD projects in the program can be funded by grants of the Evangelisches Studienwerk e.V. Villigst. Application forms can be ordered from:

Evangelisches Studienwerk e.V.  
Promotionsförderung  
Iserlohner Str. 25  
D-58239 Schwerte  
Phone: +49 (2304) 755-215, Fax: +49 (2304) 755-250

For further information about the PhD program Globalization and Employment as well as the PDF-files of the discussion papers see:

http://www.globalization-and-employment.de
THE FORMAL-INFORMAL ECONOMY DUALISM
IN A RETROSPECTIVE OF ECONOMIC THOUGHT
SINCE THE 1940s

Christine Clement

University of Hohenheim
PhD-Student and Research Assistant
at the Chair of Economic Theory (520H)
of Prof. Dr. Harald Hagemann
70593 Stuttgart (Germany)
Christine_Clement@uni-hohenheim.de
Central to the scientific debate about the ‘informal sector’ and the validity of the concept used to be a twofold challenge. The crux laid not only in the objective to explain the widely visible persistence of the informal economy in developing countries, but also in the identification of its roots and the proliferation conditions to be met ex ante. The present paper aims at establishing a link between the theories on informality and marginalization which is another important issue that has arisen within the discussions on the causes of persistent poverty a few years ago. Both concepts are interlinked and self-enforcing. On the macroeconomic level, any economy – be it formal or informal – consists of a set of different economic sectors and any of these sectors basically consists of an accumulation of people on the microeconomic level. Every time one looks at the macro level where political and economic conditions frame the dynamics of the formal and the informal economy, one has at the same time to look at the micro-level where the social and economic conditions determine the incentives for every actor to participate either in the formal, the informal or in both economies. Informality has multiple sources depending on whether the agent took a voluntary choice or had to involuntary opt-out from an institutional system. In this paper, the connection between informality and involuntary exclusion shall be examined in a retrospective of economic thought since the 1940s. The roots of the intertwined concepts of informality and economic exclusion have been laid in the dual economy theories of the 1940s-1950s. Recapitulating the works of Julius Boeke, Arthur Lewis, John Harris & Michael Todaro, Albert Hirschman and other socio-economists of that time, it will be argued that one of the necessary reasons for the persistence of the informal economy in developing countries is the dualism in institutional frameworks that leads to the marginalization of social groups and their subsequent exclusion from formal economic activities. By referring to the groundbreaking Africa studies of Keith Hart (1971) and the International Labour Organization (1972), special emphasis will be given to the causal reciprocity between informality, marginalization and economic exclusion. The paper closes with a brief overview of current schools of thought that deal very differently with the issue of informality and economic exclusion.

**JEL classification:** B20, B25, J64, O15, O17, O43, N90, P16

**Keywords:** Economic dualism, informal sector, informal economy, informality, marginalization, economic exclusion, involuntary exclusion, institutions, inequality, traditional sector, urban rural sector, stages of development, Julius Boeke

* The author would like to thank Prof. Dr. Hagemann, Prof. Arena, Prof. Lapidus and the participants of the 17th Summer School on the History of Economic Thought, Economic Philosophy and Economic History (Zaragoza/Spain, September 2014) for useful comments and suggestions.
“One does not have to spend much time in developing countries to observe how their economies are a mish-mash, combining the productive with the unproductive, the First world with the Third.”

Dani RODRIK (2011), §1

I. FRAMEWORK

Central to the scientific debate about the ‘informal sector’ and the validity of the concept used to be a twofold challenge. The crux laid not only in the objective to explain the widely visible persistence of the informal economy in developing countries, but it laid also in the identification of its roots and the proliferation conditions to be met ex ante. Still, this paper shall not be perceived as just another review of how the sometimes controversially discussed concept of the ‘informal economy’ has evolved throughout the economics literature in the last decades.
The present paper rather aims at establishing a link between the theories on informality and another issue that has arisen within the discussions on the causes of persistent poverty a few years ago, namely the issue of marginalization. HANAGAN (2008) defines marginality, or more specific, the fact of being marginalized, as a process where individuals are systematically “relegated to the side-lines of political debate, social negotiation and economic bargaining”. Both concepts, the informal economy and the marginalization of people, are interlinked and self-enforcing. In the context of informality it is therefore necessary to look at the macroeconomic conditions of markets and the microeconomic conditions of people at the same time.

On the macroeconomic level, any economy – be it formal or informal – consists of a set of different economic sectors and any of these sectors basically consists of an accumulation of people – the economic actors – on the microeconomic level. Consequently, every time one looks at the macro level where political and economic conditions frame the dynamics of the formal and the informal economy, one has at the same time to look at the micro-level where the social and economic conditions determine the incentives for every actor to participate either in the formal, the informal or in both economies. If, on the one hand, specific businesses are executed outside the formal flows of resources, goods, services and capital, these individual informal actors are, on the other hand, marginalized and economically excluded. Informality therefore has multiple facetted aspects and basically arises out of two different sources depending on whether the agent took a voluntary choice or had to involuntary opt-out from an institutional system [CHEN (2012)].

In this paper, the latter case, namely the connection between informality and involuntary exclusion shall be examined in a retrospective of economic thought since the 1940s. In the course of the present work, a twofold argument shall be developed whose structure is as follows. The first section “Roots of the intertwined issue” is dedicated to the dual economy theories of the 1940s-1950s where have been laid the roots of the intertwined concepts of informality and economic exclusion. Recapitulating the works of BOEKE (1953) and other socio-economists of that time, it will be argued that one of the necessary reasons for the persistence of the informal economy in least developed and developing countries is the dualism in institutional frameworks that leads to the marginalization of social groups and their subsequent exclusion from formal economic activities. The second section “Informality as an
issue of and for development” reviews the advent of the concept of informality in the scientific literature. By referring to the groundbreaking Africa studies of HART (1971) and the ILO (1972), special emphasis will be given to the causal reciprocity between informality, marginalization and economic exclusion. The third and last section “Recent findings and diverging approaches” eventually provides a brief overview of current classifications of the schools of thought that deal very differently and from distinctive starting points with the issue of informality and economic exclusion.

II. ROOTS OF THE INTERTWINED ISSUE

The vivid variety of informal employment and businesses that are common attributes to all least-developed and developing countries [henceforth: LDDCs]1 are usually subsumed and simplistically referred to as the ‘informal sector’. Yet, the term ‘sector’ is misleading when it comes to the heterogeneous cluster of existing informal activities that constitute the vibrant economic reality in most LDDCs. Small-scale and mostly family-owned businesses, where all kinds of products are sold in street shops, vendors’ trays and portable booths, dominate the picture of the urban areas in the less industrialized regions in the world. Although those countries which have the highest proportion of multi-dimensionally poor people2 differ markedly in their cultures and institutions, their informal economies and the people living therein share some basic commonalities. Being labor-intensive and equipped with a low level of technical support and mechanization, all these informal occupations fill multiple gaps in the satisfaction of basic needs and daily challenges: street doctors, moneylenders, vendors of

1 Since there is still no commonly agreed definition of developing countries and thus no universally adopted terminology, the author decided to combine the two most widely used terms by the World Bank and the United Nations. By merging both terms, the author describes the ensemble of those countries worldwide (two third of all countries) that have the lowest socioeconomic indicators compared to the other countries in the world – irrespective of whether they show small or continuous progresses in some of the indicators of the socioeconomic panel (developing and emerging countries, as defined by the International Monetary Fund World Economic Outlook 2014) or whether they continue to linger in a status quo of poverty, human resource weakness and vulnerability (48 least developed countries, as defined by the United Nations Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States).

2 The global Multidimensional Poverty Index (henceforth: MPI) was created as an international measure of acute poverty. Since 2010, the MPI has been included in the annual Human Development Reports of the United Nations Human Development Program). Covering more than 100 developing countries, the index complements the widely used traditional income-based poverty measures. Calculated out of 10 different weighted subindicators, the global MPI captures various dimensions of individual deprivation such as education, health and living standards. If someone is deprived in a third or more of ten subindicators, the MPI identifies them as multi-dimensionally poor. The extent or intensity to which they are poor is thus measured by the number of deprivations these individuals experience. Cf. Alkire et al. (2015)]
information and communication technologies, traders for car spare parts, garment producers, pottery-makers, open-air laundry tubs, smoking backyard fire-recycling facilities and garbage-collectors. Although some of these businesses have immobile production sites or factories, they are still not bounded to a specific building. If necessary, for example in the case of a government intervention or the forced clearing of an informal city-quarter, these businesses could instantly and easily move within little time. However, most of the informal businesses actually are inherently mobile and easy-to-carry and as such neither fixed to a special place nor dependent on specific infrastructure. The dissociation from permanent factories and concrete-built selling shops is characteristic for any kind of informal production and service provision.

Moreover, all of the informal economies worldwide are highly segmented with respect to their various economic sectors, the workplaces as well as the ethnicity and gender of their workforce. Although they comprise hybrid forms of unregistered work, criminal activities and tax evasion, the informal economy per se also provides multiple types of easy-to-hire employment and often represents the last resort (or the only possibility ever known) in terms of rudimentary social insurance for those who lack legal and social protection [CHEN (2012)]. Apart from the questions related to the emergence of informal economic structures, development economists around the world are even more concerned when it comes to the role of the informal economy for the development process of a country and its overall economic growth. For most of the researchers dealing with informality, there are direct linkages to persistent poverty [CHEN ET AL. (2005), SAMAL (2008)], for others [e.g. DE SOTO (1989, 2000)] informality is directly linked to economic exclusion.

To understand the issue of informality within its retrospective context among the various economic theories of stagnation, growth and welfare, one has to return to the beginnings of development and the so-called ‘modernization’ theory. According to ROSTOW (1956, 1960), development per se follows a path of five stages. It begins in the first stage with the traditional society. In the second stage the preconditions for take-off are developed and growth sets in i.e. the socio-political institutions change and cultural values slowly adapt to the changing economic system. The third phase is then the so-called take-off phase where the “old blocks and resistances to steady growth are finally overcome”3. The traditional society finally transcends to the stage of industrialization with the emergence and

expansion of leading industrial sectors, a high rate of investment and deep structural changes in infrastructure and economic policy; these developments represent the basis for the fourth phase. During the latter, modern technology spreads throughout the economy and the system experiences sustained but fluctuating progress. This drive to maturity then culminates in the fifth and last stage of modernization where the economy develops the capacity to efficiently adopt the most advanced technology in its production processes and thereby to expand its product matrix.

ROSTOW (1960) concluded that societies having attained this very last stage in the 20th century showed two attributes. First, real income per capita rose to an extent where it exceeded the coverage of basic needs like food, clothing and shelter. And second, drawing from CLARK (1940) and FOURASTIE (1949), the structure of the workforce changed so that as the share of the primary sector in the provision of employment declined tremendously. Right after its publication, ROSTOW’s stages-theory was criticized for the inherent assumption of stringent linearity in the development process of a country. Amongst others, GRIMM (1979), ITAGAKI (1963), KUZNETS (1965) blamed ROSTOW for not explaining how economies could move from one stage to the next and for not detailing the necessary and sufficient conditions to be satisfied for a country to be able to attain a specific stage. Although unsatisfactory on the point of empirical evidence, ROSTOW’s descriptive generalizations nonetheless found application in the American programs of development aid of the 1960s and the early 1970s [ENGGERMAN (2003)].

With the categorization of LDDCs according to their attained stages of capitalist development, ROSTOW’s theory of modernization became part of the development ideology of the Western world. Especially during his service time for the Kennedy- and the Johnson-administration in the United States, ROSTOW’s work represented a form of guiding map that channelled foreign direct investments into developing countries. The aim of these policies was to launch the poor countries’ “take-off into sustained growth” of the capitalist type [LATHAM (1998)] and to help them to move forward into a more advanced stage on the rostowian ladder of development.

4 In the third edition of his book *The stages of economic growth*, Rostow emphasized the existence of a sixth stage, namely post-maturity. In this stage, especially western societies chose to allocate more resources to social welfare and social-security nets, and resources were increasingly directed to the provision of services on a mass basis. Cf. Ibid. (1990), p. 11

To Rostow, the process of development was a linear and irreversible transition into only one possible direction. Based on this framework, he argued that the industrialized states of today had pursued a traverse-process where they went through the various stages of development one after the other over a period of several decades since the take-off of the Industrial Revolution [Rostow (1956)]. The so-called underdeveloped countries, however, got stuck somewhere along the way of modernization. They had left the first stage of the traditional society, but had not yet attained the critical stage where capital accumulates through the mobilization of private and public savings, investments expand industrial capacity and new techniques spread in the modern and the agricultural sector. In the rostowian framework of modernization, production structurally changes from labor-intensive to capital-intensive techniques when the modern sector expands and the traditional sector starts to shrink in terms of their shares in the production of overall national income. Applying Rostow’s theory to the LDDCs of the 20th and 21st century however reveals that their process of development and structural change does not fit the theory of stages. If one nonetheless sticks to the framework, one has to argue that the LDDCs had experienced, and some of them still do face, two or more of the rostowian stages at the same time.

In his works, Rostow neglected the possibility of social dualism in the sense of a temporal coalescence of two stages of development. Nevertheless, numerous of Rostow’s scientific contemporaries emphasized the existence of a social and economic dualism in the underdeveloped parts of the world. What the LDDCs experienced was different from the usual pattern of structural change where the industrial sector increases to the detriment of a decreasing agricultural sector. Most of these countries had a century-long colonial heritage and had gained their independence only in the mid-1940s and 1950s. Their socioeconomic systems had been coined by their colonial rulers and as such also their economic structures with infrastructure networks, production facilities and the degree of technology penetration. These inherited institutions also determined how equally the benefits of these structures were shared among the indigenous population and the colonists i.e. and how much trickled down through the social classes.
III. **The Theories of Social and Economic Dualism**

The idea of a dualistic development that splits society into participating agents and non-participating agents – those who are left-behind by economic progress – stems from the Dual Economy theories of the 1940s [KANBUR (2013)]. Their commonly acknowledged theoretical founder was Dutch sociologist and economist Dr. Julius BOEKE (1942, 1946) 1953). Although not well known in the economics community nowadays, BOEKE’s extensive research and detailed descriptions of the colonial society in the Dutch East Indies between 1910-1929 have delivered two essential insights which are still valid today: i) institutions differ across cultures and cannot easily be implemented elsewhere in a top-down manner, and ii) if a specific institutional framework is imposed onto another culture, the framework will only be partially adopted by the respective population; the inborn institutions will continue to prevail i.e. they remain valid for some parts of the population. This dualism of prevailing institutions then leads to divergent economic subsystems within society. With these important conclusions, Boeke had laid the foundations for numerous succeeding publications on the theory of social and economic duality such as in chronological order ROSENSTEIN-RODAN (1943), NURKSE (1953), LEWIS (1954, 1956), ORNATI (1955), HIGGINS (1956), HIRSCHMAN (1957, 1958), JORGENSEN (1961), RANIS & FEI (1961), GEERTZ (1963), DIXIT (1970), HARRIS & TODARO (1970) and SINGER (1970).

Although BOEKE’s name rarely appears in the literature today, his ideas have at least been implicitly rediscovered in the economics literature. In the *Biografisch Woordenboek van Nederland* JAQUET (1979) shortly describes the life and works of this often forgotten economist who actually had so much implicit influence on the research of the informal sector as well as on institutions and their role for development. Born in 1884 into a Baptist priest family, BOEKE later studied literature science and political science in Amsterdam. In his doctoral thesis on the critical evaluation of the tropical colonial political economy (*Tropisch-koloniale staathuishoudkunde: het problem*), BOEKE rejected the applicability of the western economic and political institutions onto the East-Indian societies. He highlighted the need for a distinct economic science for the Dutch-Indonesian colonies where the indigenous population

---

6 Boeke had revised and expanded his two earlier studies *The Structure of the Netherlands Indian Economy* (1942) and *The Evolution of the Netherlands Indies Economy* (1946) in order to publish them conjointly in his Opus Summum *Economics and Economic Policy of Dual Societies as exemplified by Indonesia* in 1953.

7 The *Biografisch Woordenboek van Nederland* is a Dutch biographical dictionary consisting of 6 volumes which were published between 1979 and 2009. The dictionary contains short biographies of well-known or notable Dutch people.
responded otherwise than expected to the economic incentives set by the western colonial institutions. Shortly after the publication of his doctoral thesis in 1910, BOEKE left the Netherlands to work as civil servant and teacher for politics and political economy at the Gymnasium William III in Batavia. In 1914 he became advisor for the Volskreditwezen; an institution that was founded as part of the Dutch imperialistic policies in to improve the living situation of the indigenous population through target programs in education, health care and rural credit cooperatives. Throughout his time in Indonesia, BOEKE continued to do research on the social and economic living conditions of the Indonesian population. After a renewed call, BOEKE finally accepted the chair of tropical-colonial political economy (later: eastern economy) at Leiden University. In 1929 he returns to the Netherlands to work as a professor of eastern Economics with lectures on the “Dualistic Economy”. Later, BOEKE’s human and scientific integrity brought him into severe conflict with the German occupiers under the NS-Regime in the Netherlands from 1940 to 1945. At his own risk and own expenses, BOEKE published the brochure Nationaalsocialistische Staathuishoudkunde in 1941 where he heavily criticized and opposed the Nationaal-Socialistische Beweging, the Dutch political party which collaborated with the German occupiers. As a consequence, BOEKE was interned in the concentration camp Buchenwald from October 1941 until September 1944 – a detention which he survived. In the postwar period, BOEKE worked again in Indonesia but had to resign and to return to the Netherlands in September 1955 due to the late sequelae of a car accident; soon after his retirement followed his death in January 1956.

Three years before his death, his opus summum, the Economics and Economic Policy of Dual Societies – As exemplified by Indonesia was published in 1953. It contained BOEKE’S collected observations of his time in Indonesia. The emphasis throughout his works had always been on the baseline argument that every attempt to directly apply western economic doctrines and theories in the less industrialized societies was inevitably doomed to fail. BOEKE emphasized that the apparent contradiction between the western and eastern systems necessitated a distinct economic policy approach taking also into account the issue of “Societal Dualism”. He did not limit the definition of dualism to a specific dimension but continually emphasized its validity for all spheres of life and its multifaceted aspects: “it finds expression in other departments of life besides the economic – in legislation and government, in law and judicature, in social
organization, but also in men’s conception of what they need, in their evaluation of things, in work and recreation, in religion and morals”.  

Basically, BOEKE (1953) reasoned that the implementation of western capitalism in East-Asian pre-capitalist countries in the course of 19th century colonization had led to a continuous disintegration of societal structures in the respective colonies. The consequence was a persistent incongruence between the imposed socio-economic settings of the western world and the originally prevailing social and economic environment in Dutch Indonesia. In his work, BOEKE generalized9 the perceived Indonesian-European incongruence for other tropical Asian territories where European colonizers had implemented western political and economic systems in eastern aboriginal societies. The Dutch-Indies thereby served as prime example of underdevelopment because BOEKE typified their economic problems for the ensemble of all the countries in the world that had experienced decades of colonial policy imposed by the “West”: “The economic problems of Indonesia are typical for a large and important part of the world, that, therefore, an analysis of these problems may […] serve as a guide to the host of inexperienced planners for the well-being of that part of the world which has not yet conformed to their western ideals”.  

In BOEKE’s work, the West referred to the industrialized European and North American countries of the 1940-1950s. The latter were in a state of high-capitalism with a “multiplicity of organizations”, i.e. they had huge industrial sectors with a multitude of small and big competitive companies, which inevitably entailed an increasing division of labor and an economic leitmotif of efficiency. BOEKE regarded corporate enterprises as the pillars of commercial exchanges. Mechanized industries with mass production and country-wide transportation networks nourished the unlimited wants of the households. In contrast to the western principles, BOEKE defined and described the eastern society to be based on communitarianism and organically determined social ties. No sharp distinction existed between economic and non-economic motives as the joint-family was engaged in self-provision; the kin parallely constituted the basic unit of production as well as of consumption. This intimate

---

8 Boeke, J. (1953), pp. 8 & 14
9 Boeke’s generalization of his Dutch-Indonesian impressions to other colonial and developing economies was subject to diverse criticism. Central to the critique were the obvious discrepancies between Boeke’s portrayal of the Indonesian economic structure and the respective descriptions of other researchers. Cf. Higgins (1956), Breman (1976a-b, 1998), Grimm (1979) and Datta (1981).
10 Boeke, J. (1953), foreword
The commingling of survival and commercial aspects was submitted to the informal institutional framework of culture and tradition that dominated the precapitalistic society. The focus was put on the production of goods; commodity-driven commercial exchanges were complementary and remained somehow outside the daily struggle for survival. For BOEKE, the basic difference between the western and eastern systems therefore laid in the difference of the specific role of the individual for society. Production in the non-interdependent local units was based on a weak division of labor and was not influenced by prices and markets but by the needs of the kin: “Not exchange but rather self-maintenance is the basis of existence; individual self-sufficiency is the dominant idea, the unit being the family”\textsuperscript{11}. Low geographic capital such as underdeveloped transportation networks, lacking infrastructure and region-wide lacking access to electricity and potable water amplified the impact of geographical isolation to the detriment of united interregional markets. The geographically bounded local markets constituted distinct economic spheres and thus led to a scattered national economy of independent markets. As the societies’ needs mainly depended on the families’ ability to produce, economic exchange generally occurred within the limits of these local markets. Based on his observations BOEKE concluded that the eastern economic order substantially differed from the intertwined capitalist principals of western capitalism because the same economic motive that drove western commerce appeared to be naturally absent in the East.

As another crucial difference between the western and native types of production, BOEKE mentioned the so-called ‘business concerns’ which lay on different foundations. The western agents and production units relied on capital and were thus able to sustain capacity enlargements as well as shocks of diminishing revenue or even the necessity of down-sizing in times of shrinking demand and economic recessions. On the other side, the native industries were much less capital intensive and were based on the use of unpaid family labor; the only response to economic shocks therefore could be to off-set parts of the labor force. But as most of the employed laborers were usually family members that depended on the joint-output, the offsetting revealed to be difficult or just socially impossible. While the western industries were directly linked to the world market, the local small businesses were excluded from interregional or international commerce and had to rely on the local markets and their local chain of suppliers, middlemen and traders. As a consequence, the contact between these two inherently

\textsuperscript{11} Boeke, J. (1953), p. 40f
different social stages of commerce – i.e. the internationally included capitalistic society and the locally constrained precapitalistic one – generated endemic difficulties. The latter culminated in intra-societal divergence which, according to BOEKE, represented the prerequisite element of societal dualism: the antithesis between two different full-grown economic systems that simultaneously expand and sometimes cooperate but also diverge in terms of their underlying institutions. Generalized to the ensemble of the LDDCs there might thus be coalescence between international markets and a village economy as well as between an urban and a rural societal system – the dominant one of which revealed to be the imported western economic structure that did not totally absorb the naturally grown social system. Neither of the systems became generalized for the whole society. BOEKE’s theory of economic dualism was thus based on strictly separated forms of systems but also tried to grasp the interdependence and interrelations among them that actually breaded the conflicts. He did not deploy on the dichotomy of the capitalistic and non- or precapitalistic sector but insisted on the fact that the dichotomy was non-material and rather concerned incongruent philosophies of life.

Basically, BOEKE classified societies according to their economic systems, their technology and their dominating form of social organization. However, no society passes discretely from one system into the other; usually there always is a certain overlapping of various social systems through time. Thus, the assumption of homogeneous full grown social styles was often criticized to be unrealistic as only very few countries in the world show perfectly elaborated and pure social styles; instead their dominating social system may combine different dominant characteristics in a blending of overlapping styles: the “prevailing social system, the remains of the preceding and the beginnings of [the] future social style”12.

Over the 15 years of BOEKE’s observation from 1914 to 1929, the Dutch Indies experienced stagnant incomes per capita [BOEKE (1953), THE MADDISON-PROJECT (2013)13] because the high population growth was not compensated by a respective growth of economic production. This was mainly due to lacking increases in capital accumulation and industrial productivity. BOEKE concluded that the eastern societies were static economies in opposition to the western societies with dynamic economic growth paths. HIGGINS (1956) blamed BOEKE for concentrating on a dichotomy of statics versus dynamism when describing the East-West

---

12 Boeke, J. (1953), p. 3f
contrast. Basically, HIGGINS assumed the business decisions of the entrepreneur to be the engine of western growth paths while the eastern economies with their production inputs (natural resources, labor and capital) did not rely on innovation and continuous structural change: “Such economies bear little resemblance to the theoretical abstraction known as a ‘static’ economy, and an economy which is not growing is still a special case of economic dynamics”. However, BOEKE asserted himself that the term “static” when being applied to the eastern societies should not describe absence of growth or unaltering production matrixes. In the aim to preserve the traditional patterns of the economy and averting any disturbing element for the communitarian institutions, the issue of stationarity was linked to informal institutions and “based on tradition, aiming at rest, beset with impeding factors and hence lacking in suppleness, and relegating economy to a secondary, subordinate position”. The staticity was thus to be equalized with economic exclusion – be it voluntary or involuntary – from societal change.

In his explanations for the prevailing social dualism, BOEKE explicitly referred to the role of predominant informal institutions in the Dutch Indies such as the social values, culture and the individual attitude towards one’s own contribution to the welfare of society. He hypothesized an underlying social tenor of fatalism and resignation towards the creation of output beyond the satisfaction of the population’s subsistence needs. An increase in wages in the traditionalist sector would thus have an insignificant effect on the incentive structure of the workers and even lead to an actual decrease in the supply of labor. HIGGINS (1956) joined this argumentation and described the traditional informal institutions as persistence of a “feudal attitude towards commerce and industry” which led to individual unwillingness towards product advances. This inherent tenor also constrained entrepreneurial spirit and appetite for uncertain investments. But “development of the capitalist type cannot take place without capitalists” as the latter enable inventors to eventually come up with a more efficient allocation of resources and inputs and thus, economic growth. For HIGGINS, the situation of the LDDCs and the therein observed slow shift in the labor force and the heterogeneous structural

15 Boeke, J. (1953), p. 29
16 The distinction between formal and informal institutions was mainly coined by the works of Douglass North (1990). In 1993, the latter was awarded the Nobel Memorial Prize in Economic Sciences together with Robert Fogel “for having renewed research in economic history by applying economic theory and quantitative methods in order to explain economic and institutional change”. Cf. http://www.nobelprize.org/nobel_prizes/economic-sciences/laureates/1993/presentation-speech.html
17 Higgins, B. (1956), p. 111
change from the traditional to more industrialized sectors were due to society-inherent rigidities with respect to the adoption of new technologies and institutions. HIGGINS and BOEKE concluded that therefrom have arisen those hybrid forms of business units and transaction processes which could be observed in LDDCs all over the world in the 1950s. This very conclusion is still valid for the informal economy concept today.

Already at the time of his publications, the works of Boeke on social dualism had inspired many other contemporary economists. For example MOORE (1953) considered the mode of employment to be the central criteria of distinction within these dualistic economic structures. For MOORE, there was a sectoral identification problem because one had to differentiate between the inactive share of the population and those being economically active or productive, i.e. those entitled by a regular occupation and having the status of an employer, employee, self-employed or unpaid family worker. However, MOORE was well aware of the fundamental difficulty to define what was ‘economic’ and thus ‘economically active’. This reasoning joined the argumentation of BOEKE that eastern communitarian societies were characterized by a division of labor with respect to age, sex and intra-kinship obligations. The adoption of capitalist principles entailed a necessary reorganization of production structures and thus an increase in specialization and the division of labor. Both consequences were incongruent with the traditional structures of the eastern kins.

Another contemporary of BOEKE was ORNATI (1955) who did a study on the nature and problems of industrial labor in India. He emphasized the contrast between the East and the West in terms of differences in the prevalent forms of employment and the prevailing informal institutions with respect to the latter. To ORNATI, the term labor force revealed to be inappropriate to describe the level of economic activity or the ratio between tapped and untapped economic potential of an underdeveloped country. In 1948, the so-called Factories Act had been passed in India. This law aimed at the regulation of the employment conditions in Indian factories, mines and the communications industry. Based on the application or consecutive violation of these formal employment regulations, ORNATI divided the labor market into two categories i.e. the organized and the unorganized sector. Yet, the majority of the Indian workers were not employed in these industries but in agricultural plantations, trade services or in the uncountable small manufacturing enterprises producing basic consumption commodities – in short, all the other occupations which were outside the rule of the Factories
Act and thus part of the unorganized sector. In this respect, ORNATI also highlighted the infeasibility of targeted policies which was due to the critical lack of statistical information on the number of ‘unorganized’ laborers, their presumably difficult or even miserable working conditions and their earnings. The aforementioned informal institutions also played a vital role when ORNATI described the very pronounced dislike for factory work of the Indian workers. The Indian informal institutions and attitudes apparently did not comply with the westernized working conditions and especially the strict regulations and timing in the industrial sector: “Occasionally, the worker leaves the factory not to return to the village but to rebel against being forced into what might be called the ‘factory norms’: time discipline, the limitation on leisure, the confines of the machines, the toil of learning, and the like”\(^{18}\). In 1947, the issue of strong ties to the rural birthplaces and the inner pressure to return (even though no job opportunity or other stake in the village economy may lie ahead), was officially labelled as Village Nexus\(^{19}\) by the Indian Government: “The worker stays in the factory until illness, nostalgia, boredom, or unemployment drives him back to the village where he stays for a week or a month until the need to earn a living returns him to the city for employment”\(^{20}\).

As already elaborated previously, the daily reality of how the poor eastern masses worked and traded was incongruent with the western capitalistic principals. The imposed institutional settings responded to the requirements of a capitalistic system with individual economic agents rather than kinship communities. Due to the seemingly incongruent economic institutions and habits, BOEKE was convinced that the majority of the eastern population was and, if the institutional status quo was to stay unaltered, would remain unable to adopt the western capitalistic structures. Some of BOEKE’s contemporaries were also critical towards the concept of social dualism. They emphasized the lack of a sound baseline model and regarded his theory to consist of nothing more than a detailed and sometimes even inconsistent description of the social conditions in the Dutch Indies from the 1910-1920s. HIGGINS (1956) subsumed that, because the eastern society was totally different from the European social order, BOEKE precipitately drew the fatalist conclusion that “economic and social development on western lines is impossible for the East” and did not focus on the issue that some existing eastern “social and cultural institutions […] may actually be] constituting a barrier to economic

---

\(^{18}\) Ornati, O (1955), p. 46f


\(^{20}\) Ornati, O (1955), p. 46f
development” as they obstructed the adoption of western economic principles.\textsuperscript{21} In response to BOEKE’s deficiencies on the institutional question, HIGGINS (1956) described various sociological barriers to economic development. Amongst them, he highlighted the organization of society around the kinship ties which may represent a disincentive for the individuals to supply more labor, save more or incur higher investment risks. The economic outcome of any individual engagement would always positively or negatively affect the kin as a whole, because the joint-family represented the basic unit of production and consumption in the eastern society.

The steady intrusion of western structures into a precapitalistic country in terms of capital, technology, organization and institutions created a gradual and continuing shift in the balance between the native production system and the western style supply chain. As for the case of the Dutch Indies, the handicraft methods’ share in overall production shrank as the traditional producers became intermediate suppliers to the nascent industries in the urban centers. Although capitalist structures connected and intermingled with the traditional sector, the population groups in the rural areas were disconnected from the capitalist enclaves in the urban areas whose structural and institutional “boundaries […] constantly become more clearly defined and more unsurmountable”.\textsuperscript{22} Thus, already BOEKE described the issue of marginalization and subsequent economic exclusion of the rural population.

Structural change per se implies not only the emergence of (temporary) dualism, but also the aggravation of inequality among the population depending on who is engaged in which economic sector i.e. in the traditional or in the emerging capitalist sector. For Dutch-Indonesia BOEKE concluded that the elite aligned its efforts and objectives with the promising western power. The cultural and monetary inflows that accompanied the instauration of western capitalism diverted their devotion (irrespective of whether the elite was part of an extractive or inclusive i.e. private property based institutional framework) from their own society towards the new economic lodestar\textsuperscript{23}. For ROSTOW (1956, 1960) economic progress arises out of a stimulus

\textsuperscript{21} Higgins, B. (1956), p. 111
\textsuperscript{22} Boeke, J. (1953), p. 215
\textsuperscript{23} In this respect, Boeke points out to the misuse of political de jure and de facto power to design institutional settings that benefit a specific clientele such as the political elite of a country. This issue had been further elaborated decades later in the field of institutional economics by, amongst others, North & Weingast (1989), North (1990), North et al. (2009) and was later subsumed under the term “extractive institutions” which was coined by Acemoglu et al. (2005, 2009, 2014). Yet, the latter think of a different direction in the link of causation of political reorientation, as they assume that the political institutions at time t influence the economic institutions at time t. The latter in turn determine the distribution of
that incites a change in the balance of powers, the social values character of the prevailing economic institutions. Such a stimulus – may it be a political revolution or a technical innovation – alters the propensities of individuals and thus spurs the economic initiative of specific social groups, e.g. the social élite, while the rest of society remains in the status quo: “More often than not the economic motives for seeking economic progress converge with some non-economic motive, such as the desire for increased social power and prestige, national pride, political ambition and so on”\(^\text{24}\). This reasoning was also later shared by Breman (1976a-b)\(^\text{25}\): “In country after country, those in control of the instruments of economic power have consistently perverted one of the essential purposes of government – to redress the major imbalances in the society – by using the power of public office to further entrench their privileged position”\(^\text{26}\). However, the mobility to switch from one sector to the other and thus the respective opportunity set was limited by the individual’s social standing and the institutions that individual adhered to and was subjected to. Based on Breman’s explanations one can argue that with the exception of the eastern ruling elites that had quickly aligned with the colonizers and adopted the new system, the majority of the population became increasingly marginalized. With the expansion of western capitalist structures and the nascent but steadily growing exports-producing agricultural industries, the traditional businesses became increasingly economically excluded. As a consequence, the social dualism eventually manifested itself on the microeconomic level in terms of diverging individual opportunity sets between the ruling élite and the rest of the population. The latter became more and more constrained in their selectable options. For Boeke, this deprivation of the trickle-down elements of capitalism (i.e. the fruits of trade created through the increased dovetailing with international commerce) annihilated the “dynamic, developing element in the [eastern] culture”\(^\text{27}\), a loss that inevitably led to economic stagnation for the majority of the population.

To sum up: throughout his research, Boeke defined the ‘West’ and the ‘East’ as two discrete socio-economic systems that were embedded in their very own specific institutional

\(^{24}\) Rostow, W. (1956), p. 27

\(^{25}\) Breman, J. (1976a), p. 1875

He also provides a vague path towards a solution for the issue of exclusion through the gain of access to formal schooling. Cf. Breman, J. (1976b), p. 1906


\(^{27}\) Ibid. p. 39
frameworks with their distinctive culture and philosophy of life. The socio-economic settings in the western world diverged from the social and economic environment in the tropical Asian territories in such a way as both systems were ruled by very distinctive institutional orders. The same socio-economic doctrines that were applied in the western societies, were only partially valid or not at all congruent to the eastern institutional systems: “When one distinguishes all these peoples according to their degrees of culture and their political situation, one perceives that every group moves in its own economic order and that every theoretic system has to rest upon its own economic order to be usable for a given nation”.28


BOEKE’s theory of social dualism represented the theoretical background for various research projects in the 1970s on the advent and persistence of the informal economy in LDDCs. As different as the latter are with respect to their culture and history, they nevertheless share a common past as former colonies. Most of the LDDCs in Africa and South East Asia had inherited institutional dualism from their former colonial rulers and remained unable to resolve it – even after gaining independence in the 1940s and 1950s. Challenged by massive migration waves from the rural areas to the expanding cities, the prevailing institutional heterogeneity reinforced the marginalization of social groups and thus their economic exclusion from formal markets [HART (1973)].

28 Quote of an unknown Indian author, quoted in: Boeke, J. (1953), p. 8
Informality and engagement in the shadow economies became common attributes of the steadily growing cities in all developing and transitioning countries.

IV. INFORMALITY AS AN ISSUE OF DEVELOPMENT

Migration played a central role in the development models of the dualism economists of the 1950-1960s. Lewis (1954, 1956), who is generally considered to be the founder of the mathematical dualism models, was convinced that the basis of the development process lay in the flows of resources, i.e. labor, capital and goods, between the industrialized and the traditional agricultural sector. In theory, already small technological advances set off surplus labor in the structurally changing agricultural sector and improved health care induces further population growth. The rural excess labor flows to the urban areas where the industrial sector expands. The latter is assumed to produce the machines that will even more increase productivity in the agricultural sector and thus offset more surplus labor to be re-employed in the industrial plants. A parallel flow of increasing agricultural food surplus will in turn nurture the industrial workers and thus, increase their productivity. This synergy of economic flows between various sectors of production was also central to the discussion of the displacing and compensating employment effects of new technologies. These enduring controversies had been initiated by Ricardos’s (1821, 3rd edit.) book chapter On Machinery and had inspired a new strand of literature focusing on the issue of the Traverse Process\textsuperscript{29} from structural change. Especially the ‘horizontal’\textsuperscript{30} growth models such as the one from Lewis (1954) emphasized the sectoral interdependencies that occurred within an economy and modelled the intersectoral flows of production factors. In the Lewis-model, a two-sector economy with a modern and a subsistence sector is assumed. While the modern sector uses reproducible capital whose product can be sold above wage-costs, the subsistence sector does not. Output in the latter is lower than in the industrialized sector because productivity is not “fructified by capital” and thus remains “unproductive”\textsuperscript{31} i.e. below its hypothetical potential.

\textsuperscript{29} Hagemann, H & R. Scazzieri (2009), pp. 143-147

\textsuperscript{30} There are basically two different kinds of formal representations in the history of economic thought where the implications of structural change were modelled: i) the vertical models from i.a. Lowe (1976) and the ‘later’ Hicks (1973), and ii) the horizontal models from i.a. Lewis (1954, 1956) and the ‘early’ Hicks (1965).

\textsuperscript{31} Lewis, A. (1954), p. 407f
LEWIS saw the central problem in the theory of economic development in the understanding of how capital accumulation (including physical capital as well as knowledge and skills) was triggered. He tried to understand the institutional and the structural change that incited the ruling ten percent of a preindustrial country to save more than the rest of society i.e. the process by which a community with low saving rates converts itself into a growing economy where voluntary saving increases to more than twelve percent of national income. For [LEWIS (1968)] the crux was the limited financial capacity for investment among the population: capital availability therein was restricted so that funds for business purposes and trade were cheaply accessible, but not obtainable for investments in buildings and agriculture. Through the highlighting of the unequal access to financial institutions, LEWIS already considered a specific form of economic exclusion. He concluded that the financial marginalization had wide-reaching and long-term impacts on a country’s prospect for economic development. An often cited phrase of LEWIS in this respect is that “the tendency for capital to flow evenly through the economy is very weak; in a backward economy it hardly exists. Inevitably what one gets are very heavily developed patches of the economy, surrounded by economic darkness”32.

Central to the Lewis-model was also the inherent assumption of an ascending sectoral hierarchy of the traditional sector with respect to the economically growing industrialized sector. This sectoral hierarchy was also emphasized by BREMAN (1976a-b). The latter actually considered this very unequal interdependence between the two diverging sectors to be preconditional for development as it could be seen as an intra-system mechanism of resource reallocation: “The low rate of industrialization and the presence of surplus labour are listed as principal reasons why a dualistic system has sprung up in the cities of the third world. The informal sector contains the mass of the working poor whose productivity is much lower than in the modern urban sector from which most of them are excluded”.33

Basically, the Lewis-model lay on three fundamental characteristics: i) dualism of systems, ii) unlimited supply of unskilled labor at the current wage, and iii) unskilled labor is paid more in the modern sector than in the traditional sector for the same quantity and quality of work. The first aspect was inspired by BOEKE and became more systematically described in LEWIS’ later works from 1968 and 1979. The other two conjectures became the fundamental

32 Lewis, A. (1954), p. 409
assumptions in the growth models of the 1950s and 1960s. By assuming an ‘unlimited supply of labor’, LEWIS separated both sectors not only in terms of productivity and but also in terms of their employment structures: “[the supply of labor is] so large relatively to capital and natural resources, that there are large sectors of the economy where the marginal productivity of labour is negligible, zero or even negative. [This applies also to the large sector of] the whole range of casual jobs – the workers on the docks, the young men who rush forward asking to carry your bag as you appear, the jobbing gardener, and the like. These occupations usually have a multiple of the number they need, each of them earning very small sums from occasional employment; frequently their number could be halved without reducing output in this sector.”

In his revision of the Dual Economy (1979), LEWIS also highlighted the informal character of the traditional sector in LDDCs to which his model could be applied. In 1968, LEWIS published a paper where he sought to clarify and expand his previous works in several respects. With the expression ‘unlimited supply of labor’, the second fundamental model characteristic, he meant absolute elasticity of labor supply at the existing wage in the capitalist sector that continues to exceed the earnings in the agricultural sector – inciting continuous waves of migrants that leave for the urban areas. In later publications on the dual economy theory LEWIS (1968), (1979) abandoned the phrase “unlimited supply of labour” in favor of an “infinitely elastic supply of labour to the modern sector at the current wage” in order to avoid confusion and, as he often emphasized, “emotional distress”.

The third basic characteristic of the Lewis-model was the wage differential between the modern and the traditional sector. The wage level in the modern sector depends upon the earnings in the subsistence sector: the “minimum at which labour can be had is now set by the average product of the farmer; Men will not leave the family farm to seek employment if the wage is worth less than they would be able to consume if they remained at home.”

---

34 Lewis, A. (1954), p. 403
35 Lewis also clarified that his model was never to be understood as an anti-socialist attack; although he considered economic growth to arise out of industrialization, Lewis never explicitly outlined other systems than the capitalist production. Moreover, in his model, he made no difference between private or public capitalists that hired agricultural surplus labor for profit. Moreover, he strictly denied any disparagement of the peasant production vis-à-vis the capitalist sector. It was just an assumption that in the initial stages the dynamism of growth was due to in capitalist expansion. To Lewis, the agricultural sector only temporary served as a reservoir of labor (later complemented by immigration and increased participation of women) and was by no means bound to a state of zero-growth: “The model does not deny that peasants can grow rich by producing more, or more valuable output; it does not argue that capitalist production is more valuable; it is not normative” [Ibid. (1968), p. 2f]
36 Ibid. (1979), p. 211
37 Ibid. (1954), p. 409
reinvestment of the profits of industrial production within the modern sector, capital accumulation increases. The industrial sector expands and thereby drives rural-urban migration, attracting even more labor out of the traditional sector. Up to a certain extent agricultural output remains the same due to a negligible marginal productivity of labor. However, at some point, capital formation in the modern sector proceeds faster than surplus labor is set free and the population grows. Through the outflow of labor, the average product per man in the subsistence sector rises automatically. The ever fewer number of farmers have to support an ever greater number of industrial workers. The labor surplus diminishes and the expansion of the modern sector increases overall demand for food and thus prices of subsistence goods. Wages in the modern sector are driven up. Lewis concluded that this was the reason why industrial and agrarian revolutions always arose together. To him, a stagnant agricultural sector was the crucial constraint to industrial expansion and thus economic growth.

The interdependency structure between the traditional and the modern sector became also a central argument in Hirschman’s (1958) *Strategy of Economic Development*. He assumed the wage differential between both sectors to be a veritable growth advantage for the preindustrial sector. Joining the argumentation of Lewis (1954, 1956, 1968), he reasoned that the preindustrial sectors in the LDDCs would be spared by the forces of industrialization for a longer time due to their labor-intensive and less capital-intensive structures. The far lower wage in the preindustrial sector thus represented a comparative advantage for the underdeveloped economies compared to the already industrialized sectors in the western countries. The lower labor costs would enable the independent (not to say informal) small-scale manufacturers and petty-commodity traders to dominate the market for specific products where the fix costs for advanced industrial methods would simply exceed the expected profits to a large extend. Hirschman therefore saw a huge dormant economic potential in the labor-intensive production methods of the informal sector e.g. the production of furniture, shoes, apparel, ceramics, baskets, basic metalworking, food-processing as well as house construction and local transportation. All these businesses would thus not only serve as employment buffer for the less skilled (obsolete) workers but would also generate economic growth through the increase in overall income. Hirschman’s argumentation in his early work on dualism was dominated by a transitory view of dualism and optimistic expectations with respect to its impact on long-run development: “[Dualism] has some compensating advantages and represents in a way an
attempt by the economy of an underdeveloped country to make the best out of its resources during a transitional phase”.

However, in his later works on collective action and institutions, HIRSCHMAN changed his argumentation on the role of the preindustrial sector and on dualism in general. Although he did not explicitly mention the term informal sector in his writings, the ‘later’ HIRSCHMAN (1974, 1978, 1981) considered economic dualism and the continuous presence of sectoral divergence to be the result of activist reactions to discontent with the prevailing institutional framework. Basically, HIRSCHMAN (1978) differentiated between two possible reactions an individual could choose from when faced with unfavorable organizations. The individual could either stay and maintain the relation with the organization while voicing his complaints and hoping for improvement; or the individual could exit and interrupt his interrelation with the respective organization because there was no hope for future improvement. This political framework of action could be used to explain the persistence of informal structures as a result of social and economic exclusion and the therefrom marginal status of specific population groups. The exit of individuals from the formal economy may thus be viewed as a decision motivated by despair and unsurmountable barriers to personal development. HIRSCHMAN (1981) also emphasized this point when he argued that exits occur in response to “the arbitrary and capricious actions of the sovereign”. While voicing one’s complaints may be a costly process in terms of time and money, the exit-option reveals to be less costly and easier to realize as the individual may rely on social ties within its kin to assure his survival in the institutional framework of the informal sector. With this argumentation Hirschman also sought to explain the ever growing unemployed labor force that accumulated in the urban areas.

Only with HARRIS & TODARO (1968, 1970) the final differentiation between the various coalescing sectors was introduced: the rural agricultural sector and the rural urban sector where traditional technology prevailed in opposition to the equipment of the urban modern sector. They did not believe in the assumption of an agricultural labor surplus being responsible for the high unemployment rates in the urban areas and the ever-accelerating rural-urban migration.

38 Hirschman, A. (1958), p. 132

Lewis (1979) emphasized a similar line of reasoning. To him the expansion of informal small-scale activity in urban areas was conducive for development because with its products and services it satisfied genuine needs and provided employment for the migrants from the countryside.

waves. Instead, they assumed that the conventional wage and price adjustment in the labor market were annulled and could not counterbalance the magnetic attraction of the ‘bright lights’ of the cities onto the peasants. Contrary to the development models before, and based on their own empirical observations in Southeast Asia and Africa, they aimed at explaining the migration dynamics in LDDCs through the rural-urban differences in expected earnings. The urban wage level was assumed to have a dual function i.e. the determination of the industrial employment level and the allocation of labor resources between the traditional and the modern sector. Yet, the urban wage level was usually politically and not economically determined and thus represented an incentive for a continued rural-urban migration. Vis-à-vis the wage differential between the traditional and the modern sector, migration thus became an economic rational choice on the part of the job-seeking rural individuals. Therefore, all targeted subsidy-policies for reducing urban unemployment and ‘back to the land’ programs such as the ones implemented in the 1960s in Kenya and Tanzania would be ineffective. They would probably prevent the effects of the urban minimum wage on unemployment but would of course not increase the level of industrial employment [HARRIS & TODARO (1968)].

Similar to LEWIS’ argumentation, HARRIS & TODARO emphasized the inherent intersectoral connectedness through the migration of labor. There was a high elasticity of rural labor supply with respect to the creation of additional employment in the industrial sector i.e. every additional industrial job likely attracts more than one agricultural worker. Due to the limited absorption capacity of the industrial sector, the surplus migrants find themselves in a so-called urban traditional sector, the equivalent of the later concept of the informal sector. Especially TODARO (1969) stressed the role of the urban traditional sector with respect to rural-urban migration: Rural migrants behaved as absolutely rational maximizers of expected utility in terms of expected marginal urban real income. Although there was uncertainty whether the expected income would immediately exceed the present real rural income, migrants considered it to be still economically worth wile from a long-run perspective [i.e. “from a discounted present value approach to the rural-urban work choice”, TODARO (1969, p. 140)] to venture for the cities. Migration should thus be considered as a two-stage phenomenon. The informal sector builds the stage of arrival for the unskilled rural migrants and the modern sector represents the second stage where these migrants might eventually find a permanent job in the formal economy. If they fail, they remain on the first stage. Some may return to their rural home towns, others may
stay. For the latter the opportunity to work informally and gain the informal urban wage may still outweigh the opportunities awaiting them in their rural home areas.

Instead of dividing the economy into two distinct formal and informal spheres such as HARRIS & TODARO, BREMAN (1976a) emphasized the “unity and totality of the productive system” as a continuum. He considered the unequal distribution of labor opportunities in a fragmented and mutually exclusive labor market to be the root cause for economic dualism and the division of the economy into a formal and an informal segment.40

V. FROM THE INFORMAL SECTOR TO AN INFORMAL ECONOMY

Being ‘informal’ basically refers to being not registered. Although not accounted for in the national statistics, there is substantial economic activity taking place in the informal sector thus also every day in those countries that are actually said to be underdeveloped. For roughly 45 years now, the vivid variety of these non-registered activities had been simplistically referred to as ‘informal sector’. The advent of this catch-all term is commonly attributed to have originated in two independent publications, namely the work of Keith HART (1971, 1973)41 and the so-called Kenya-Report of the INTERNATIONAL LABOUR ORGANIZATION (henceforth: ILO) from 1972. At that time, the denomination ‘informal sector’ was not backed by a uniform definition but was solely based on the description of the complexity of the income and expenditure patterns and the urban economic structures in two Sub-Saharan LDDCs.

In his analysis of the living conditions of the Frafra minority in Ghana, HART (1973) described the dual reality of the urban labor market of Accra. For several decades, the south-Ghanaian harbor city had been facing a continuous inflow of laborers migrating from the rural areas and seeking for employment in the capital. The absorptive capacities of the Accra labor market did not grow proportionally and thus the increasing surplus labor accumulated in the outskirt slums of the city. Remaining unabsorbed, these migrants sought for other, eventually informal, income-generating activities besides the formally accounted businesses. HART (1973) concluded

---

40 Breman, J. (1976), pp. 1871-1874
41 Contrary to the usual citations, the original article “Informal Income Opportunities and Urban Employment in Ghana” had been written by Keith Hart in 1971. The final publications of his article, however, only occurred in 1973: i) by Jolly, R. et al. (eds., 1973) in: Third World Employment: Problems and Strategy, Penguin, Harmondsworth; ii) in the Journal of Modern African Studies, Vol. 11, No. 1, pp. 61-89, Cambridge University Press, Cambridge (UK). For simplicity, the final publication date “1973” will be used throughout the citations of Hart in this paper.
his article by emphasizing the two-sided character of the divergence between the formal and the informal structures and their universal presence throughout all LDDCs: “Accra is not unique, and a historical, cross-cultural comparison of urban economies in the development process must grant a place to the analysis of informal as well as formal structures. It is time that the language and approach of development economics took this into account”\(^\text{42}\). With his work, HART was the first author\(^\text{43}\) to establish a consensual analytical term – i.e. the informal sector – in order to unite all the uncountable appellations [such as subsistence, traditional, preindustrial or unorganized sector] that had been randomly used in the scientific discussions on the dualistic economic structures in LDDCs.

Similar conclusions were drawn in the final report of the pilot country mission of the ILO World Employment Program in Kenya from 1971-1972: “For those who cannot make a living in the rural areas or fulfill their aspirations, there is the alternative of migrating to the towns. Particularly among the younger and the better-educated men […] there is a great influx to the towns, where both the formal and the informal sector of the economy offer income-earning opportunities.”\(^\text{44}\) The informality concept had enlarged the initially narrow notion of unemployment in the development literature through the integration of a huge variety of workforms beyond formal employment. The ILO concluded that the unemployment explanation often used in the literature and based upon the Lewis-model could not explain the persistence of the informal economic structures around the cities in LDDCs. Unemployment was often analyzed as direct result of a basic imbalance within the primary and secondary sector of an economy. In times of structural change there may be a discrepancy between a rapidly growing labor force and a more slowly growing number of job opportunities. If unemployment is too large and no safety nets are available to the job seekers to assure their subsistence living, there are two hypothetical opportunities for these people to choose: they could either return to their home regions and their kin or stay in the urban areas with an inherent danger of destitution. However, in reality, people do not only stay, but they also do not immediately starve. Instead, even more people migrate to the cities. These people find a subsistence living that is more attractive than a potential return to where they came from, where there is no

\(^{42}\) Hart, K. (1973), p. 89


\(^{44}\) ILO (1972), p. 10
employment either. The informal sector absorbs all of these people and assures their survival – at least at a subsistence level.

The introduced terminology of a ‘formal’ and an ‘informal’ sector sought to more adequately describe the sectoral duality in developing countries. Instead of insisting on the bias of the traditional-modern dichotomy, the authors of the ILO-report emphasized that “both sectors are modern; both are the consequence of the urbanization that has taken place in Kenya over the last 50 years. We might have used the terms ‘large-scale’ and ‘small-scale’, but those terms are purely descriptive and tell us nothing about why one sector is large-scale and the other is small-scale”\textsuperscript{45}.

The informal sector cannot be considered a transient phenomenon that arises as a form of petty-commodity production in the course of structural change. Instead, the informal sector represents a persistent phenomenon as it can be considered as market economy in its purest form where the quantities and prices of the goods exchanged are determined by their supply and demand without any distortionary state regulation. The magnitude of participation of the agents within the informal sector, however, is determined by their voluntary or involuntary choice of economic status and may be triggered by the economic exclusion of these individuals. The informal sector therefore represents a persistent attribute to all non-communitarian societies – be they capitalist, socialist, developed, transitioning or still struggling to industrialize.

The introduction of the informal sector in the 1970s made it necessary to adapt the hitherto employed terms with respect to the issue. The usual dichotomy between the traditional/subsistence vs. modern/industrialized sector had thus to be abandoned in favor of a more adequate distinction between the formal and the informal sectors of an economy. Yet, even with a different terminology, still no straight demarcation line of affiliation could be drawn between both sectors and their agents’ occupations. The ILO-report (1972) highlighted that any differentiation between both sectors should be based on the “way of doing things”\textsuperscript{46}; the report therefore enumerated several criteria of distinction based on a theory of ‘access’ and ‘constrained opportunities’. Informal activities were thus characterized by an ease of entry, the reliance on indigenous resources; family ownership of businesses, the small-scale of operation, the absence of any formal support and legal regulation, high competitiveness of markets and

\textsuperscript{45} ILO (1972), p. 503f
\textsuperscript{46} ILO (1972), p. 6
the acquisition of skills outside the formal school system. In opposition to these aspects, activities and businesses in the formal sector were based on corporate ownership, large-scale and more capital-intensive operations, integration into international markets and reliance on resource and technology imports, protected markets and highly regulated activities through detailed legal frameworks and trade licenses, as well as formally acquired skills and knowledge. However, the most important attribute of formality per se was the difficulty of entry into that framework. As such, both, the formal economy and the informal sector were assumed to work entirely isolated from each other. In the 1970s, the informal sector was though regarded as a form of detached subsector being encapsulated within the formally accounted economy.


However, the increasingly wide-spread motivation among economists to analyze the informal sector made it crucial to back the thitherto catch-all term with a harmonized statistical definition.

Upon its 17th International Conference of Labour Statisticians in 2002, the ILO eventually adopted principal guidelines to enable the empirical grasp of the informal sector in its entirety with its various forms of employment and business activities. Thereby, the ‘informal sector’ definitorily evolved to an ‘informal economy’ that “refers to all economic activities by workers and economic units that are – in law or in practice – not covered or insufficiently covered by formal arrangements. Their activities are not included in the law, which means that they are operating outside the formal reach of the law; or they are not covered in practice, which means that – although they are operating within the formal reach of the law, the law is not applied or not enforced; or the law discourages compliance because it is inappropriate, burdensome, or imposes excessive costs”\(^47\). On the one hand, the newly adopted definition more adequately

\(^{47}\) ILO (2014), p. 4
reflected the employment environment under which the urban and rural poor [i.e. those parts of a given population with the lowest incomes] earned a subsistence living in LDDCs. The definition now also allowed for the accountancy of the size and the dynamics of the informal economy in both, developing and industrialized countries based on internationally harmonized criteria.

Broken down to its basic statutory characteristic, formality is based on the fact of being registered and being legally liable for any committed action. Through their status, agents in the formal sector therefore have a privileged access to social, legal and financial services. As already seen in the discussion of the works of BOEKE earlier in this paper, these privileges were often granted to the ruling elite of a developing country that quickly adapted and engaged with their resources in the new economic system [cf. also ACEMOGLU ET AL. (2005)]. This assumption of unequal access – be it historically given or an effect of post-colonial resource redistribution, is a fundamental issue of dualistic societies and thus central to every discussion on the formal and the informal sector. Already BOEKE (1953) and HIGGINS (1956) linked the concept of sectoral divergence to constraints on the availability of resources, and thus access to these very resources. In their Kenya-report, the ILO (1972) also highlighted the link between the economic situation of people and their possibilities of access to education, other social facilities as well as resources, and through that, also the close link to urban migration, the “frustration of unsuccessful job seekers” as well as the situation of the so-called (informal) working poor that “[work] very hard and strenuously, but their employment is not productive in the sense of earning them an income which is up to a modest minimum”.48

Based on what had been illustrated until here, it becomes obvious that the issue of dualism, in its social and economic sense, was recurringly linked to the issue of accessibility of resources and the concreteness of individual opportunities in the sense of having the choice between different alternatives concerning one’s own future path of development. Therefore, there has always been a direct connection between the dualism of the formal and the informal economy and the exclusion of certain groups within society.

48 ILO (1972), p. 9
Excluded agents find themselves on the margins of society in various respects. Against the background of poverty and development, the concept of exclusion describes a process of amplification and reinforcement. To the best of the author’s knowledge, De Soto (1989, 2000) was the first to describe an explicit link between the marginalization of people and their (subsequent) engagement in the informal economy. The exact relation between informality and economic exclusion\(^{49}\) was never clearly defined throughout his works. Nonetheless, De Soto had no doubts that the consequences therefrom were conducive to keep the majority of the population in poor countries in an endemic state of poverty. He defined marginality as a lack of means to define, benefit from or enforce one’s claim to (economic) rights, and its cogent outcome being the desperate necessity to become extra-legal in order to gain a living despite the economic status on the margins of society.\(^{50}\)

De Soto’s approach to poverty reduction was based on entitlement i.e. the procurement of titles to land and businesses to formalize people, make them accountable, officialize their collateral and thus, eventually tap their dormant economic potential. The fact of being extra-legal entailed a specific crux; although most of the poor people already possess at least some assets (be it a dwelling, a piece of land, a shop with a specific customer base), they hold no formal title to these possessions thus they dispose of as sets in defective\(^{51}\) form. In De Soto’s view this crux to development lay in inefficiently constituted legal systems that actually suppressed incentives for investment.

\(^{49}\) Initially, the term was used to describe the social fragmentation and the dissolution of social connections between the society as a whole and specific individuals, ethnicities or entire social classes. The term since then spread into the field of development economics. It is commonly assumed that the authorship of the term ‘social exclusion’ appertains to French politician Lenoir who, in 1974, wrote about the social exclusion encasing one tenth of the population in France: “To say that a person is unfitting, marginal or unsocial simply corresponds to stating that, in an industrial and urbanized society at the end of the 20th century, this very person is incapable of satisfying her needs […] or finds herself segregated – be it due to her own reason or be it the reason of collectivity” (the author’s own translation). Cf. Lenoir, R. (1974), p. 10.

\(^{50}\) De Soto, H. (2000), chapter 1

\(^{51}\) De Soto regarded capital in the form of assets as a dead value that had to be reanimated and brought into a form that could be used in economic transactions. The formal property system of the industrialized countries endows the asset holder with socially recognized entitlements to use a resource (usus), to appropriate any returns from this use (usus fructus), to change the form and substance of the resource (abusus) and to alienate these very rights to a third person (ius abutendi) to benefit from the usus, the usus fructus, abusus and ius abutendi of his asset. For a more comprehensive introduction to property rights theory, please refer to the discussions in Steiger, O. (2007).

Analogously to Demsetz’ (1967) explanations, De Soto’s development approach was though not about the disposition of assets with specific technical characteristics, but about bundles of specific rights that were defined and backed through a particular set of legal institutions and that determine the socially restricted use of commodities. Cf. Demsetz, H. (1967), p. 347f
These systems failed to provide legally backed representations of assets in formal property documents. The latter would enable the respective asset holders to create additional value and maybe even increasing returns to their accumulated capital.\textsuperscript{52}

The aforementioned definition of marginality by DE SOTO was further developed by HANAGAN (2008) in his contribution to the \textit{International Encyclopedia of the Social Sciences}. VON BRAUN \& GATZWIELER (2014) finally set up a definitory framework of marginality and exclusion to fill the semantic vacuum opened up by DE SOTO’s use of these terms. Marginality was thus defined as “the position of people on the edges, preventing their access to resources and opportunities, freedom of choices, and the development of personal capabilities”. The denomination of poor individuals to be socially or economically excluded from progresses in health, technology, infrastructure, business opportunities, or public services in any form indicates that these individuals are relegated to the margins of social and economic life. Exclusion thus inevitably leads to marginalization.

VON BRAUN \& GATZWIELER (2014) conclude that marginality is also a root cause of poverty because it describes the involuntary position of population groups at the margins of the political and economic system. The legal and geographical disablement of access to resources, public services and legal enforcement of the peoples’ rights necessarily cause an unequal distribution of opportunities. This description of dualism in access rejoins the aforementioned lines of reasoning of Boeke. Through the influence on the individuals’ disposable capability sets\textsuperscript{53}, social and economic dualism have long term effects on the economic performance of society as a whole.\textsuperscript{54} HANAGAN (2008) emphasized the significant overlapping between different forms of exclusion as “groups that are excluded in one area tend to be also excluded in other areas”\textsuperscript{55}. The same tenor was propounded by the World Bank (2003): “While different national discourses vary as to which aspect of exclusion – economic, social, and/or political – they emphasize, there is general consensus that the inability of individuals or social groups to

\begin{itemize}
  \item \textsuperscript{52} De Soto, H. (2000), chapter 1
  \item \textsuperscript{53} The theory of the capability approach was developed by Amartya Sen [cf. i.a. Sen (1985, 1999)]. The approach is based on the individual freedom to choose given the choice. The disposable variety of selectable alternatives arises out of individually accessible assets, rights, knowledge and opportunities. Capability sets are thus determined within the formal and informal institutional framework surrounding the individual within a society.
  \item \textsuperscript{54} Von Braun, J. \& F. Gatzweiler (eds.) (2014), p. 3
  \item \textsuperscript{55} Hanagan, M. (2008), 1st paragraph
\end{itemize}
participate fully in the economy, in social life, and in political processes reduces social solidarity, augments social tensions, and holds back social development”.

VII. RECENT FINDINGS AND DIVERGING APPROACHES

The dual development of societies with persisting dualistic institutional frameworks may explain why, in DE SOTO’s words ”people who have adapted every other Western invention, from the paper clip to the nuclear reactor, have not been able to produce sufficient capital to make their domestic capitalism work”57. As it has been elaborated throughout the preceding chapters, many theories have made their advent when it came to the explanation of the evolution and the persistence of the informal economy. In order to bring more structure into the multiplicity of overlapping and partly opposing approaches, CHEN (2012) grouped the different strands of theory on the emergence of the informal sector into four different schools.

First, the Dualist School identifies only limited interconnections between the formal and the informal economy. Informal agents are regarded to be excluded from modern business perspectives and employment opportunities because of a discrepancy between the population growth rates and the increases in actual industrial employment as well as an incongruity between the skills supplied by the labor force and the abilities and know-how demanded and required by modern industries. Second, the Structuralist School views the informal engagement to be driven by the structural nature of capitalism and economic growth per se. In an attempt to circumvent labor costs and enhance competitiveness, formal firms optimize their economic situation by reallocating their labor force into informal employment. Third, the Legalist School, founded by the works of De Soto (1989, 2000), argues that those people engaged in informal activities try to circumvent a legal system that is perceived to be inappropriate by bypassing legal regulations with informal laws, rules and norms. The formal institutional framework is counterproductive to the growth of small and medium-sized enterprises and thus hampers incentives to accumulate capital. Fourth, the Voluntarists reason that the decision to work in the informal sector is taken deliberately after weighing the benefits, and especially the costs, of

56 World Bank (2003), p. 18
57 De Soto (2000), p. 7; this phrase also became the winged word of De Soto’s ‘Mystery of Capital’.
leaving the formal sphere or of not entering the formal economic system but of remaining in the shadow economy.

When it comes to the interrelation of the informal economy with a country’s development process, the scientific opinions differ once more strongly on the issue. More recently LA PORTA & SHLEIFER (2014) summarized the different research strands into three different and in some points opposing schools of thought. These are concerned with the character of reciprocity and interrelation between the formal and informal economic structures. This categorization has been taken up and further expanded by the author of this paper and will be briefly described hereafter.

The first category is represented by the legalist and property rights economists like DEMSETZ (1967) and DE SOTO (1989, 2000). Especially the latter focused on burdensome government regulations, unreliable legal protection and the lack of formal titles to property. These deficiencies spur disincentives to register businesses and to trust the perceived defective formal institutional framework. DE SOTO assumed that informal microbusinesses could potentially be as productive as formal businesses if only their untapped potential could be freed. For him, the crux lies in reforms to the property registries and the lowering of administrative barriers to the processes of business registration. With formal titles to their property, as little as the latter may be, the property holders would be able to obtain credit from the formal banking system and could thus invest in their businesses, accumulate capital and expand – thereby fueling and spreading economic growth. In DE SOTO’s approach the formalization of informal structures neutralizes the status of being on the margins of society. The exclusion of formal property titles impedes the access to credit funds, saving accounts and insurances and, thus, economically marginalizes the small entrepreneurs and drives them into the shadow economy. DE SOTO attributed these differences in access to institutional settings that favored the societal elites and disfavored the rest of the population. Recapitulating the history of Peru and the important legislative changes with respect to the Peruvian informal traders, DE SOTO (1989) concluded that the informal economy had to be considered to be both, traditional and institutionalized.58

---

58 In his first book “The Other Path”, De Soto traced back the historical legislative evolution of informal trade in Peru. After the first documented regulation of informal trade of street vendors in 1594, several other laws on the issue had been passed in the subsequent decades and centuries. Depending on the respective predominant political tenor, there had been alternating attempts to regulate, repress, formalize and even legally support informal street trade. [Cf. De Soto (1989), pp. 75-79]
The institutional disproportions in rights and access led to an extension of the informal sector because they limited the structural absorption of migrants venturing to the cities in search for wage jobs. The hence involuntarily excluded migrants had to rely on family connections and thereby entered an informal network of supply, processing tasks and trade which mainly takes place in domestic workshops and dwelling-shops. De Soto’s legalist approach can thus be considered to be rooted in the dual economy theories of the 1940s-1950s and the approach of Harris & Todaro of the 1970s.

In opposition to the ‘legalist’ approach, the second category comprises those economists that highlight the aspect of illegality and the potential “parasitic” character of the informal sectors with respect to the formal economy [La Porta & Shleifer (2008), p. 109f]. In 2004, an influential study of the McKinsey Global Institute (henceforth: MGI) was published that pointed out the problems of the persistent informal sector in transitioning countries such as Turkey, Poland and the BRICs. Within these countries, the informal businesses and companies were and still are in direct competition with the formally registered businesses. For the last three decades all of these countries have attracted huge amounts of foreign direct investment in manufacturing. Vis-à-vis the continuous inflow of funds, the informal enterprises fostered and preserved their artificial competitive advantages that increased their productivity. Through the avoidance of taxes, social security contributions and the non-compliance with labor regulations these firms artificially lowered their overall cost burden: “MGI found that the substantial cost advantage that informal companies gain by avoiding taxes and regulations more than offsets their low productivity and small scale. Competition is therefore distorted because inefficient informal players stay in business and prevent more productive, formal companies from gaining market share” [Farrel (2004), p. 28].

According to the MGI-study, the often praised short-term employment buffer function of the informal economy is altered to a long-term growth obstruction and welfare setback because the creation of new jobs within the formal sectors is hampered. Moreover, wages do not rise but are kept low – not to say bidden down – through the oversupply of cheap informal labor which, in addition to that, usually is not insured and not registered with social security. Farrel summarized two central economic consequences of informality. First, the prevalence of companies within the informal economy retains them from realizing economies of scale and gains in productivity. Second, the informal companies that evade taxation and labor legislation
are thereby enabled to steel market shares from their formal competitors that comply with the rules. However, FARREL failed to substantiate these findings with sound statistical evidence instead of anecdotal estimations for the countries in question.\textsuperscript{59} Another major default of the MGI study is that it blindly generalized results from the BRICs countries to the whole developing world.\textsuperscript{60}

One central fact that is not contested across the literature on informality is that the informal economies of industrialized, transitioning and LDDCs are very heterogeneous in their structure. The sectoral weights i.e. the focus of activities within the informal economy are different from country to country. The diversity ranges from subsistence survival, self-employment and unpaid professional work to tax evasion, drug cartels, mafia networks, prostitution rings, and illegal gambling operations. Considering the criminal aspects of the informal economy, the second category of economists – who may be called ‘extralegalists’ – were in favor of suppressing informal businesses instead of condoning or even supporting their informal status.

Besides the aforementioned \textit{Extralegalists}, the third category consists of those authors who focus on structural dualism and market inefficiencies that lead to the advent of an unofficial shadow submarket within the formal market economy. Among these \textit{Institutional Dualists} are HARRIS \& TODARO (1970), LEWIS (1954, 1956), MAZUMDAR (1976), ROSENSTEIN-RODAN (1943) and ROSTOW (1960). They considered the traditional sector (i.e. the equivalent of the later termed informal sector) to be a transient and ephemeral attribute along the development path of an economy from the traditional handicraft economy to the industrialized mass production. Incentives to enter the informal economy are created through labor market restrictions for immigrants, trade barriers for merchants and, on the individual level, the gap between the total cost of labor and the net earnings after tax deduction i.e. the tax wedge which basically represent foregone earnings.

Others within the group of the \textit{Institutional Dualists} considered informality to be a natural side effect of poverty in times of structural change where the traditional sectors and the modernizing sectors of an economy diverge. BANERJEE \& NEWMAN (1993) regard the process

\textsuperscript{59} In Russia, the MGI-estimated disadvantage to supermarkets caused by price dumping of informal food retailers is about 13 \% of the final good’s price. In Turkey, the cost savings of informal dairy processors is estimated at almost 20 percent compared to their formal law compliant competitors. Cf. Farrell, D. (2004), p. 32

\textsuperscript{60} Ibid., p. 32
of development as a path-dependent and self-enforcing dynamic process that depends on the initial distribution of wealth and the occupational choices taken thereafter. They built a model to illustrate the interplay between the initial distribution of wealth and market imperfections which determine the choice of occupation of the economic agents and the process of development of that very economy. Those classes of society among which certain resources are concentrated have more freedom to choose among different opportunities and may become employees, self-employed or remain unemployed. Moreover, market inefficiencies in the capital market lead to very unequal supply of funds to the various classes of society. Constraint in borrowing and sometimes even deprived from any financial services, poor people are unable to choose business occupations which require critical amounts of investment to be started or to be run. Instead of becoming entrepreneurs themselves, these people choose, or are forced to choose, to work as dependent employees with a wage rate that ensures the clearing of the labor market. “The pattern of occupational choice is therefore determined by the initial distribution of wealth, but the structure of occupational choice in turn determines how much people save and what risks they bear”. The occupational choices of the individuals depend on the initial distribution of wealth and the thereby facilitated or constrained possibilities of accessing credit. The model inherent capital-constraintness of individuals emphasizes the possibility of increasing returns to capital-access or, as Banerjee and Newman termed it, “pecuniary increasing returns”.

With the works of Romer (1986, 1990, 1994) and Lucas (1988) and the emergence of endogenous growth theory in the 1980s and 1990s, the focus was put on increasing returns to scale induced by technological progress. The development economists of these years, however, focused on market imperfections and the therefrom resulting constraints and unequal access to public services and financial market services. In their dynamic non-linear model of occupational choice, Banerjee & Newman (1993) endogenized the distribution of capital and assets within society in order to highlight the path-dependence of the labor market structure and subsequent overall economic development. There is either convergence to a high and wealthy steady-state or to a low equilibrium of low employment and low wages where parts of the population are excluded from the whole range of possible occupational choices.

---

62 Ibid., p. 277
The previously mentioned *Structuralists* [La Porta & Shleifer (2008), Levy (2008), Ordóñez (2010), Schneider (2012) and Schneider & Enste (2002)] focused on structural – not to say institutional – inefficiencies that are created through the absence of legal frameworks with respect to wage legislation and labor security. Overall, these economists agree that the dynamics of the informal economy are closely linked to the inability of governments to universally enforce their authority in terms of supervision of tax compliance and the payment of social security contributions. In the same sense, Rauch (1991) focused on the effects of government’s ability to enforce wage laws where large firms were usually monitored but small firms were neglected. The latter were then able to pay less than the formal economy’s “wage floor” to their employees and to hire them on a day-to-day basis. Theoretically, the unequal enforcement of labor and wage regulations between larger (formal) firms and small (informal) businesses would lead to biased competitiveness and productivity, and thus, ostensibly, to biased market prices. Informal entrepreneurs are investment-constrained because they lack access to most public goods and financial market services; thus they diverge from formal entrepreneurs with respect to their production structures and have to rely on the use of low-skilled human capital and low-quality inputs. For Rauch informality, in some sense, is thus also an inevitable byproduct of poverty. As a direct consequence, the allocation of human and natural resources between the formal and the informal economy is determined through the ability to buy resources and to employ labor. The gap in productivity due to difference in the endowment of resources also causes the emergence of a price- and quality-determined dualism in labor forms and business sizes.

However, La Porta & Shleifer (2008) reject the idea that simply registering informal firms would eliminate the gap in productivity between formal and informal firms. They argue that this very gap arises out of the divergence in the production structure. There is a genuine labor market dualism as the labor wages and goods prices informal firms deal with generally reflect the differences in productivity and technical sophistication. An imposed formalization through the subjection to legal enforcement would destroy the informal market economy and cause unemployment of millions of people that hardly achieve their subsistence-living.

This last chapter on the various schools of thought dealing with the issue has shown that the concept of ‘informality’ has been evolving for the last 60 years. The contemporary employment

---

Rauch, J. (1991), p. 34f
of the concept is that of an economic continuum where a distinct informal subeconomy is encapsulated within the formally accounted economy. At its advent, the concept of economic dualism described the coexistence of a traditional subsistence economy vis-à-vis a modern industrialized sector. The perception of the developing economy as a dual economy with one part being integrated into the world economy and another part remaining outside is considered to be the first examination of what is today known as the issue of the formal-informal sector continuum. Boeke’s analysis of social dualism and his generalization of this coexistence to other LDDCs where western culture had penetrated eastern lands marked thus the beginning of the research of the informal economy. Boeke (1953) regarded the issue of dualism in the developing world as almost unsurmountable and therefore as persistent attribute. However, his theoretical successors of dualism viewed it as a transitory phenomenon arising due to a different phasing of development. Lewis (1954, 1956) and Harris & Todaro (1970) considered a growth framework where rural surplus labor flowed into the industrialized sector and expanded the theoretical production capacity. However, as these constant inflows were not met by respective increases in the availability of resources or resource productivity, economic growth lagged behind the structural changes. As a consequence, the practical production capacities were unable to absorb the agricultural surplus workers.

The last decades of research in the developing world have shown that the economic dualism between the industrialized, i.e. formal, and the still traditional, i.e. informal, sectors in LDDCs did not disappear. At the high times of the informal sector theories in the 1970s, informality immediately became a controversially discussed issue in the ongoing conflict between the capitalist and communist blocs [Hart (1973)]. The socialists regarded the informal sector as the inevitable consequence of foreign capitalist dominance in post-colonial economies where the majority of the urban population was condemned to deprivation and exploitation. On the other side, the liberals regarded the informal sector to be a temporary buffer for 'bootstrap' operations that would lift the underdeveloped economies out of their misère through their own dynamic entrepreneurs. As always, the truth lies somewhere in the middle. Although initially viewed as a natural transitory phenomenon, development research in the last 60 years revealed that the informal economy was indeed not a residue from a preindustrial phase of development but a permanent attribute especially of LDDCs where the majority of the population continues to earn their living in informal occupations.
In an attempt to describe the issue of economic parallelism, the terms formal and informal economy represent the denominators of two distinct economies, which are at different stages of development but coexist at the same time within a given country. These two economies are intertwined in a form of systemic continuum where the informal subeconomy is separately encapsulated within what is called the ‘formally accounted’ economy. This continuum of economies is based on the temporal overlapping of two divergent socio-economic settings within a specific region. Both systems, the formal and the informal economy, concur and overlap in certain sectors, but they are not congruent with respect to their underlying institutional systems.

VIII. CONCLUDING REMARKS

In the aftermath of World War II, development policies of the capitalistic countries (as opposed to the Soviet bloc) were motivated by the successful recovery of western Europe and the continent’s, especially Germany’s, experience in socio-economic growth. At that time, the unsuccessful development programs of the World Bank and the IMF in the South America, Africa and Asia were motivated by the political containment and the belief in trickle-down effects of the implementation of liberal growth policies and the attraction of foreign direct investment.

The present working paper attempted to make a clearing contribution with regard to the roots of the concept and the role of economic exclusion therein. Special emphasis has therefore been attributed to the assessment of the interrelations between economic exclusion and the naissance of the informal sector in the development economics literature. Institutionally determined marginalization of specific population groups – usually the poor who did not dispose of de facto political power – entails the latter’s deprivation of resources and opportunities. The economically excluded groups exit the formal framework of the modernizing industrial sector to gain their living in the informal sector. Informality and being an agent in the informal sector is a double-sided sword. On the one hand, the informal sector provides a free-to-access market economy where resources and labor are purely priced and exchanged with respect to supply and demand. All the transactions are only determined by the prevailing informal institutional setting. On the other hand, the therein prevailing incentive scheme tendentiously reinforces the prevailing dualistic reality of economic opportunities for the population. Based on the
explanations given in this paper, the author concludes that one necessary reason for the advent of the informal economy is the marginalization of social groups and their subsequent exclusion from formal economic activities. Moreover, there is a two-way causal relationship between the engagement in the informal economy and the fact of being marginalized and economically excluded. This causal reciprocity incites a self-enforcing cycle that leads to a status quo of informality and economic exclusion where the latter is partly voluntarily chosen and partly system-inherently forced upon certain social groups.

BIBLIOGRAPHY


<table>
<thead>
<tr>
<th>Nr.</th>
<th>Jahr</th>
<th>Autor(INNEN)</th>
<th>Titel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/1998</td>
<td></td>
<td>Bernhard Holwegler und Hans-Michael Trautwein</td>
<td>Beschäftigungswirkungen der Internationalisierung, eine Studie aus- und einfließender Direktinvestmenten der Metall- und Elektroindustrie im Raum Stuttgart</td>
</tr>
<tr>
<td>2/1998</td>
<td></td>
<td>Heinz-Peter Spahn</td>
<td>Heterogeneous Labour, the Unemployment Equilibrium, and the Natural Rate</td>
</tr>
<tr>
<td>3/1998</td>
<td></td>
<td>Philip Arestis, Iris Biefang-Frisanco Mariscal and Harald Hagemann</td>
<td>Capital Shortage Unemployment in Germany and the UK</td>
</tr>
<tr>
<td>5/1999</td>
<td></td>
<td>Hagen Krämer</td>
<td>Dienstleistungen: Motor für Wachstum und Beschäftigung in Deutschland?</td>
</tr>
<tr>
<td>6/1999</td>
<td></td>
<td>Jürgen Kromphardt</td>
<td>Lohnbildung und Beschäftigung</td>
</tr>
<tr>
<td>7/1999</td>
<td></td>
<td>Ewald Walterskirchen</td>
<td>Beschäftigungspolitik in Österreich</td>
</tr>
<tr>
<td>8/1999</td>
<td></td>
<td>Reiner Franke</td>
<td>Lohnzurückhaltung, Beschäftigung und (zu) einfache empirische Zusammenhänge</td>
</tr>
<tr>
<td>9/1999</td>
<td></td>
<td>Peter Kalmbach</td>
<td>Implications of Integration for Wage Formation and Employment</td>
</tr>
<tr>
<td>10/2000</td>
<td></td>
<td>Arne Heise</td>
<td>Theoretische Grundlagen einer Verhaltensabstimmung der makroökonomischen Politikträger</td>
</tr>
<tr>
<td>11/2000</td>
<td></td>
<td>Eckhard Hein und Carsten Ochsen</td>
<td>Monetary Interest Rates, Income Shares, and Investment: Theory and Empirical Evidence for France, Germany, the UK, and the USA</td>
</tr>
<tr>
<td>12/2000</td>
<td></td>
<td>Guntram R. M. Hepperle</td>
<td>Airbus – ein gelungenes Beispiel für beschäftigungsorientierte Industriepolitik?</td>
</tr>
<tr>
<td>13/2000</td>
<td></td>
<td>Bernhard Holwegler</td>
<td>Implikationen der Technologiediffusion für technologische Arbeitslosigkeit</td>
</tr>
<tr>
<td>14/2000</td>
<td></td>
<td>Markus Schreyer</td>
<td>Wachstum und Beschäftigung vor dem Hintergrund des Solowschen Produktivitätsparadoxons</td>
</tr>
<tr>
<td>15/2000</td>
<td></td>
<td>Mauro Boianovsky</td>
<td>Some Cambridge Reactions to The General Theory: David Champernowne and Joan Robinson on Full Employment</td>
</tr>
<tr>
<td>16/2001</td>
<td></td>
<td>Marc-Peter Radke</td>
<td>Law and Economics of Microsoft vs. U.S. Department of Justice: A New Economic Paradigm for Antitrust in Network Markets or Inefficient Lock-in of Antitrust Policy?</td>
</tr>
<tr>
<td>17/2001</td>
<td></td>
<td>Daniel Hartmann</td>
<td>Taylor-Regel und amerikanische Geldpolitik</td>
</tr>
<tr>
<td>Nr.</td>
<td>Year</td>
<td>Author(s)</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>----------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>18/2002</td>
<td>2002</td>
<td>Jutta Maute</td>
<td>Stabilization via Currency Board</td>
</tr>
<tr>
<td>19/2002</td>
<td>2002</td>
<td>Daniel Hartmann</td>
<td>The Fed Strategy: Successful but Out-of-Date?</td>
</tr>
<tr>
<td>24/2007</td>
<td>2007</td>
<td>Ralf Rukwid</td>
<td>Arbeitslosigkeit und Lohnsprizeung - Empirische Befunde zur Arbeitsmarktsituation gering Qualifizierter in Deutschland</td>
</tr>
<tr>
<td>26/2008</td>
<td>2008</td>
<td>Andreja Benković &amp; Juan Felipe Mejía</td>
<td>Tourism as a Driver of Economic Development: The Colombian Experience</td>
</tr>
<tr>
<td>28/2009</td>
<td>2009</td>
<td>Constanze Dobler</td>
<td>The Impact of Institutions, Culture, and Religion on Per Capita Income</td>
</tr>
<tr>
<td>30/2009</td>
<td>2009</td>
<td>Patricia Hofmann</td>
<td>Die neue neue Außenhandelstheorie: das Melitz-Modell</td>
</tr>
<tr>
<td>31/2010</td>
<td>2010</td>
<td>Julian P. Christ</td>
<td>The Geography and Co-location of European Technology-specific Co-inventorship Networks</td>
</tr>
<tr>
<td>32/2010</td>
<td>2010</td>
<td>Julian P. Christ</td>
<td>Geographic Concentration and Spatial Inequality: Two Decades of EPO Patenting at the Level of European Micro Regions</td>
</tr>
<tr>
<td>33/2011</td>
<td>2011</td>
<td>Finn Marten Körner</td>
<td>An Equilibrium Model of ‘Global Imbalances’ Revisited</td>
</tr>
<tr>
<td>36/2012</td>
<td>2012</td>
<td>Georg Erber und Harald Hagemann</td>
<td>Wachstums- und Investitionsdynamik in Deutschland</td>
</tr>
<tr>
<td>37/2012</td>
<td>2012</td>
<td>Ralf Rukwid</td>
<td>Grenzen der Bildungsexpansion? Ausbildungsinadäquate Beschäftigung von Ausbildungs- und Hochschulabsolventen in Deutschland</td>
</tr>
<tr>
<td>38/2012</td>
<td>2012</td>
<td>Vadim Kufenko</td>
<td>Empirical Analysis of Regional Economic Performance in Russia: Human Capital Perspective</td>
</tr>
<tr>
<td>Nr.</td>
<td>Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>39/2013</td>
<td>Harald Hagemann und Vadim Kufenko</td>
<td>The Political Kuznets Curve for Russia: Income Inequality, Rent Seeking Regional Elites and Empirical Determinants of Protests during 2011/2012</td>
<td></td>
</tr>
<tr>
<td>40/2014</td>
<td>Danila Raskov und Vadim Kufenko</td>
<td>The Role of Old Believers’ Enterprises: Evidence from the Nineteenth Century Moscow Textile Industry</td>
<td></td>
</tr>
<tr>
<td>41/2014</td>
<td>Harald Hagemann, Georg Erber, Niels Geiger, Johannes Schwarzer und Oliver Zwiessler</td>
<td>Wachstums- und Investitionsdynamik in Deutschland</td>
<td></td>
</tr>
<tr>
<td>43/2015</td>
<td>Christine Clement</td>
<td>The Formal-Informal Economy Dualism in a Retrospective of Economic Thought since the 1940s</td>
<td></td>
</tr>
</tbody>
</table>

Die Betreuung und Vernetzung der Arbeiten im Promotionsschwerpunkt entspricht der Arbeitsweise eines Graduiertenkollegs. Betreuer des Schwerpunkts und Herausgeber der Schriftenreihe sind folgende Hochschullehrer:

Prof. Dr. Harald Hagemann                                  harald.hagemann@uni-hohenheim.de
Institut für Volkswirtschaftslehre (520H)
Universität Hohenheim
D-70593 Stuttgart

Prof. Dr. Heinz-Peter Spahn                                  peter.spahn@uni-hohenheim.de
Institut für Volkswirtschaftslehre (520A)
Universität Hohenheim
D-70593 Stuttgart

Prof. Dr. Hans-Michael Trautwein                             michael.trautwein@uni-oldenburg.de
Fakultät II – Institut für Volkswirtschaftslehre
Carl von Ossietzky Universität
D-26111 Oldenburg

Dissertationsprojekte im Schwerpunkt können durch Promotionsstipendien des Evangelischen Studienwerks e.V. Villigst gefördert werden. Bewerbungsunterlagen können unter folgender Adresse angefordert werden:

Evangelisches Studienwerk e.V.
Promotionsförderung
Iserlohner Str. 25
58239 Schwerte
Tel.: 02304/755-215, Fax: 02304/755-250

Weitere Informationen zum Promotionsschwerpunkt sowie die Diskussionsbeiträge der Violetten Reihe im PDF-Format finden sich im Internet unter der Adresse:

http://www.globalization-and-employment.de