



TI 2009–105/3

Tinbergen Institute Discussion Paper

Is More Entrepreneurship better?

Philipp Koellinger¹

Christian Roessler²

¹ *Erasmus University Rotterdam, and Tinbergen Institute;*

² *Dept. of Economics, Brown University, USA.*

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam

Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

Tinbergen Institute Rotterdam

Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at
<http://www.tinbergen.nl>.

Is More Entrepreneurship Better?

Philipp Koellinger (1) (2) (3)

Christian Roessler (1) (4)

Submitted version (v1.1)

Last update 03.12.2009

Abstract

We develop a new perspective on the boundary of the firm that is consistent with the empirical observation that the share of entrepreneurs first decreases and then increases in the course of economic development. Existing theory based on transaction costs (Coase, 1937, and subsequent literature) is difficult to relate to these well-established dynamics. Our approach focuses on changing incentives to specialize and adapt, in order to access complementarities that arise from diverse abilities and wealth. We discuss why the efficient number of entrepreneurs is bounded and changes over the course of economic development.

Keywords: Entrepreneurship, theory of the firm, organizations, economic development

JEL Codes: L26, D20, J24

(1) Erasmus University Rotterdam, Department of Applied Economics, P.O. Box 1738, 3000 DR Rotterdam, The Netherlands. Email: koellinger@ese.eur.nl, tel.: +31-10-4081625.

(2) Tinbergen Institute, Rotterdam, The Netherlands.

(3) Erasmus Research Institute of Management, Rotterdam, The Netherlands.

(4) Brown University, Department of Economics, Robinson Hall, 64 Waterman St, Providence, RI 02912. Email: christian_roessler@brown.edu, tel.: +1-401-863-2179.

1. Introduction

Taking advantage of pockets of individual ability in society (i.e. specialization) was a life-or-death matter for the early hominids, after repeated climatic shifts about 30,000 years ago encumbered their ways of life and made efficient hunting and food gathering critical to survival. While the Neandertals failed in this and diminished until they could no longer sustain themselves (Finlayson and Carrión 2007), the ancestors of the modern humans survived in these conditions because of a number of relative advantages. These included division of labor and long-distance exchanges (Wong 2009). Among Neandertals, women and children joined the hunts, which reduced the efficiency of the food supply and exposed entire communities to dangers (Kuhn and Stiner 2006). Although Neandertals, like our human ancestors, cooperated, they were not sufficiently productive in the absence of specialization. They lacked, in the fundamental sense of the word, entrepreneurs in their midst.

In the modern human economy, productivity improvement is often cited as the recipe for sustained growth, and many would agree with Schumpeter and Kirzner that such improvements are largely carried out by entrepreneurs – perhaps the more the better.¹ Empirically, however, the fraction of entrepreneurs in a population declines in the early stages of development and trends up in mature economies. In other words, for long stretches in economic development, there is a negative correlation between growth and the share of entrepreneurs in a population that reverses only in highly developed economies. Apparently, there exist dynamic constraints that drive either the efficient level of entrepreneurship or systematic deviations from efficiency. Unfortunately, standard theories of the boundary of the firm offer little help in understanding the empirical pattern of firm sizes through time. We propose an alternative theory that explains changing firm size as a response to changes in exogenous conditions and the wealth distribution. Within this framework, we can consider the question of the efficient level of entrepreneurship.

¹ Although neither Schumpeter nor Kirzner explicitly argued that more entrepreneurship is unconditionally desirable, their analyses illustrated the importance of entrepreneurs to the process of economic development (Schumpeter) and the adjustment of markets toward equilibrium (Kirzner). Many policy discussions seem to be based on the notion that, because entrepreneurs have these critical economic functions, more entrepreneurship is always a good thing for the economy.

We argue that the optimal number of entrepreneurs and the optimal size of firms are ultimately constrained by the distribution of talent, on the one hand, and the state of institutions (e.g. capital markets, property rights) and technologies (e.g. economies of scale) on the other. We take abilities and the evolution of institutions and inventions as given, but we treat the productivity of feasible technologies as an outcome of economic choices. The catch-all phrase “economies of scale” refers, from our perspective, to low coordination costs, which reflect that workers have made complementary investments to a technology that is operated by an entrepreneur. Taken as an exogenous factor, economies of scale have no predictive power: it remains unexplained why technologies with such characteristics come into existence in a specific historic era, and why this trend did not continue indefinitely. We suggest that differences in talent and material resources create (and, over the course of development, alter) the scope for complementarities and incentives to collaborate under common management.

The dynamics of our approach are driven by three processes. First, exogenous improvements in infrastructure tend to reduce the costs that arise in managing workers in a firm due to communication barriers, principal-agent problems and so forth. Advances in transportation, communication and legal conditions readily come to mind. These developments make it easier for people to work together in organizations and benefit from complementarities with the resources and skills of an entrepreneur.

Second, people have a choice between specializing (mastering a particular technology and acquiring entrepreneurial skills) and adapting (acquiring worker skills that increase the potential fit with someone else’s technology and ideas). A greater share of specialists in the population increases heterogeneity and thereby increases management costs, because cognitive constraints make it difficult to communicate effectively with many different individuals.

Third, we propose that the changing distribution of wealth had predictable implications for the incentive to specialize or adapt. Total wealth not only increases through economic development, but also tends to diffuse through society via the mechanism of increasing wages. With concentrated wealth, there are fewer business opportunities and stronger incentives to adapt in order to benefit from the talent and resources of the most able entrepreneurs. With spread-out

wealth, there are more business opportunities and stronger incentives to specialize in order to exploit them effectively.

The optimal number of entrepreneurs is at any point in time determined by the specialization and adaptation choices, which reflect the evolving access to wealth and the availability of infrastructure and know-how. Individual talent determines who specializes and who adapts. Those at the bottom of the talent pool benefit by accessing the ingenuity of the best entrepreneurs in a fashion we formally analyze in Roessler and Koellinger (2009).

More entrepreneurship is not always better because specialization comes at the expense of foreclosing complementarities between people with different talents and resources. The optimal share of entrepreneurs tends to decrease in early development up to a point where wealth is widely distributed and the benefits of specialization outweigh the benefits of adaptation for a growing fraction of the population. This fraction will not grow forever, as long as the talent distribution remains essentially the same. Even in developed economies, it is efficient for many individuals to forego entrepreneurship and instead adapt to the productive ideas of others.

2. The U-Shape of Entrepreneurial Activity

For the longest time in human history, small production units dominated, and principal-agent relationships were often the result of slavery or serfdom rather than occupational choice. As shown by Finley (1973: p. 65), “historically speaking, the institution of wage-labour is a sophisticated latecomer” (cited in Wennekers 2006: p. 22). The available empirical evidence on the prevalence of self-employment dates back to the early 19th century in some countries and clearly indicates a U-shaped relationship with economic development. Since the industrial revolution, there was a decline in business ownership and self-employment from around 50% in 1800 to around 10% in the 1970s, and a corresponding rise of wage employment, in Western Europe, Japan and North America (Wennekers et al. 2010, Wennekers 2006, Steinmetz and Wright 1989, Edvinsson 2005, Phillips 1962, Storey 1994). Hence, firms became larger, and the small business sector seemed to evaporate. This development was driven by a decline of the

share of agricultural employment as well as a decline in the non-agricultural self-employment rate.

A reversal can now be observed in the most highly developed economies, including the US and Germany. The self-employment rate in the US began to rise in the early 1970s (Fain 1980), and this trend was mirrored in most OECD countries in the period between 1972 and 2007 (Wennekers et al. 2010). Evidence for a U-shaped relationship between entrepreneurship and economic development was explicitly presented by Acs, Audretsch and Evans (1994), Carree et al. (2002, 2007), Freytag and Thurik (2007), Van Stel and Carree (2004) and Wennekers et al. (2010), using different panel data and empirical methods. Hence, there is now robust empirical evidence that the century-long decline in self-employment has stopped and actually reversed in most high-income countries.

A similar picture emerges in cross-sectional studies that compare self-employment levels across countries: the least developed countries typically exhibit very large shares of self-employment and new firm creation. This share is much lower in developed countries and typically increases in the most developed ones (Clark 1960, Kuznets 1971, Yamada 1996, Loutfi 1992 and Bosma et al. 2008).

3. Individuality and Incentives to Specialize vs. Adapt

The U-shape can be regarded as one of the stylized facts in the history of the firm that needs to be reflected in the theory of the firm. Typically, economies of scales that are linked to the invention of particular technologies are advanced as an explanation of these patterns. But economies of scale are themselves outcomes of individual choices that simplify human interaction. They are unlikely to exist if people find it costly to commute, do not speak a common language, and cannot commit to agree to work for the same wage every day.² By learning languages and conventions, acquiring literacy and relocating to population centers,

² Acemoglu (1998, 2002) presents a related argument yielding incentives to invest into technologies that are either complementary to low or high skilled workers.

people overcome some of the coordination costs that must be borne in agency relationships. Since these are necessary ingredients in collaborating under one roof and taking advantage of investments in machinery and systems, economies of scale are not merely a by-product of technological inventions or a fact imposed by the timing of discoveries. They reflect the population's capacity for coordinated work.

It would not have been possible to run Adam Smith's pin factory without a supply of workers that spoke a common language, adhered to common behavioral standards and shared some basic reading and arithmetic skills that wage contracting required. The primary question is therefore: what led individuals to homogenize in the early stages of industrialization, and what leads them to acquire more heterogeneous sets of skills and characteristics today? These choices – the subject of this paper – combine with technological developments to give rise to greater or smaller economies of scale.

A trade-off arises from the diverse abilities individuals possess. Heterogeneity makes coordination more difficult, but it also allows us to be more productive by collaborating under the guidance of the most able. The latter aspect entails more than the benefits from division of labor and has rather different implications: constraints to personal ability can be relaxed not only by focusing on narrow instead of broad tasks (specialize), but also by accessing other people's ideas and skills within hierarchical organizations (adapt).

In the economy, there are at any point in time those who specialize and those who adapt, although their relative numbers may change. A specializer explores the productive possibilities his or her particular environment (location, culture, social standing, personal resources) offers and aims to become maximally effective at something. This can, but does not have to, mean narrowing one's skills – some technologies involve more flexibility, others less. The distinctive quality of specializers is that they develop expert skills that are particularly valuable in the context of their own characteristics and circumstances. They “represent” a technology – a set of capabilities that are suitable for creating value in their environments.

The alternative to specialization, i.e. adaptation, involves acquiring skills that facilitate learning and applying technologies. Someone who has chosen to adapt instead of specialize can efficiently collaborate with others, carry out tasks they devise with minimal loss due to

miscommunication and error. This strategy allows individuals who (due to ability, resource or personality constraints) lack the capacity to develop a valuable technology of their own, to connect with those who do have it.

It is instructive to consider how our perspective relates to that of Lazear (2004, 2005), who predicted that entrepreneurs will be “generalists,” and workers will be “specialists.” The generalist entrepreneurs have the big picture in mind: how to combine tasks, carried out by specialized workers, into a valuable whole. We do not deny that the entrepreneur needs to have sufficiently broad knowledge to coordinate operations. When we speak of specialization, we refer to a technology (i.e. a bundle of tasks), not a single task. We differ in nuances on the role of workers. They may, of course, have a particular skill, but they must be able to function within the world of the entrepreneur, which the technology and language of the firm necessarily references. In fact, employees are usually selected for “fit” with the organization, and they have a strong incentive to invest in a general ability to fit.

Hart and Moore (2002) raise the related question whether someone with “coordinating” or “specialist” capacities will have authority in a hierarchy. “Coordinator” types have ideas, if any, for a set of multiple assets. “Specialist” types have ideas, if any, for the use of a single asset. Authority is here defined as a priority of ideas: if two individuals have an idea for using an asset, whose idea is implemented? Ideas involving multiple assets optimally have priority over ideas involving fewer assets. Hence a “coordinator” is optimally assigned authority, provided that the likelihood of having an idea diminishes in the number of assets covered. A “specialist” works under the authority of a coordinator who covers the specialist’s asset. If the synergies between assets are large (i.e. ideas often involve many assets jointly), firms tend to be large; if synergies are small, firms tend to be small and independent.

Even though our entrepreneurs “specialize” in a technology, they fulfill the role of Hart and Moore’s coordinators in integrating and directing the activities of the firm. Synergy between assets is a difficult variable to observe that can perhaps be viewed as determined by the adaptation choices of workers. In this case, more adaptation leads to larger firms in our theory and theirs. Thus, while some of the literature seems to have emphasized “opposites” of

specialization as entrepreneurial qualities, contradictions with our approach stem primarily from terminology and a focus on different functions of the entrepreneur.

The heterogeneity of individuals, which creates potential complementarities but also coordination costs in working together, implies that the efficient fractions of entrepreneurs and employees in the population are not zero. In Roessler and Koellinger (2009), we derive the optimal identities and number of entrepreneurs from values that agents can generate under each other's management. These values may reflect specialization and adaptation decisions. Efficient organizations form from individual choices between self- and wage-employment that take advantage of the unique abilities of some individuals by maximizing complementarities.

The effects of individuality on the optimal number of entrepreneurs are highlighted by the extreme cases. If all agents are identical (same talents, resources and access to technologies), then there are no gains from giving up independence and collaborating with others, forming organizations and introducing management structures. At best, collaboration in joint organizations will yield the same returns as when all agents act as self-employed entrepreneurs. And if there are any management costs, the returns to collaboration will be strictly negative. Clearly, in such a situation the optimal arrangement is that everyone is an entrepreneur and all firms are one-agent enterprises.

However, when there are any differences between individuals in some economically relevant dimension, there is scope for complementarities: opportunities for two individuals to create more value together than independently. Few would be willing to share knowledge, ideas, experience, capital and other personal assets without a mechanism that allows for appropriation of some of the gains. Organizational integration and direct management is such a mechanism. By hiring others, talented individuals can put their ideas and assets to greater use and generate additional income for themselves and their workers, who would otherwise not have access to these productive resources.

Then why do we not all work in one firm, with the most talented entrepreneur of our time at the top? One factor that limits integration are (physical and verbal) communication barriers. For example, the fact that early hominids lived in fragmented populations that were geographically widely dispersed (Fabre et al 2009) meant that the specialized talents of the best hunters and

gatherers could not be accessed by many others. One can speculate that some tribes would have benefitted from management by a more able hunter or gatherer from another tribe. But collaboration across tribes was largely precluded by the costs of overcoming geographical, cultural and language barriers that made communication and coordination difficult, as well as by lingering territorial or other conflicts. As a result, people collaborated in small groups and followed the lead of the best hunters and gatherers in the tribe.

When geographic and communication barriers are overcome, we usually observe more, but not complete, integration. Today, separate firms operate within close proximity of each other: in the same building, perhaps on the same floor. These firms are specialized and require different kinds of entrepreneurial talent. One individual is unable to be a master of all things: entrepreneurs in modern economics operate unique production processes and need to possess enough specialized knowledge to direct them productively. Diverse experiences and languages (arising from culture or professional specialization) make it more difficult for people to manage each other effectively. Hence the extent to which people with different characteristics can efficiently work together is limited.

One kind of cost associated with specialization is language specificity, which Cremer, Garicano and Prat (2007) study. Smaller, more focused, organizations can adopt a more precise language than larger organizations; this is a relative cost of running a larger organization, especially in a high-tech environment. Efficient languages have precise words for frequent events and vague words for rare events. Endogenously, language becomes more precise when certain events occur often in the organization (the distribution of events is “unequal”). Heterogeneously skilled members (with limited capacity to process information) will develop separate languages (in the Cremer et al. model, “dialects” are never optimal) and can therefore not communicate. While superiors in an internal hierarchy can function as “translators” (i.e. substitute for horizontal coordination between departments), organizations are in general subject to “decreasing returns to scope”: the need to create a common language or invest in translation limits the diversity of services performed by the firm.

4. A Dynamic Theory of the Firm

Up to now, we have discussed how firm size and the level of entrepreneurship can be explained by the effects of specialization and adaptation choices on management costs. Our purpose is to enable a dynamic theory of the firm. This step is necessary if we wish to consider whether “more entrepreneurship is better” with reference to the empirical fact that entrepreneurship, for a long time, was in decline (and later increased). An argument from mechanisms that cannot accommodate the U-shape would be speculative.

How does the process of economic development affect the trade-off between the gains and costs of specialization in a way that seems to increasingly favor the creation of large firms and fewer entrepreneurs, and then later on a reversal of this trend? We suggest that two trends led to this pattern. First, technological progress, improvements in physical, legal and institutional infrastructure as well as increasing levels of general education reduced management costs and thereby softened restrictions on the extent to which people can work effectively under common management and benefit from complementarities.

This process led, second, to increasing overall wealth and growing income for large parts of the population. Wealth was highly concentrated initially, when there was little of it, and increasingly dispersed as total wealth grew. For some individuals, it is beneficial to access the resources of others through an employment relation, but this requires effective communication, physical commuting, complementary skills etc. (in short, a homogenization between individuals).

Dispersed wealth creates demand for specialized products that can only be supplied efficiently with advanced technology that favors skill specialization. It takes more specialized (hence, more) entrepreneurs to run the firms that exploit these opportunities. As wealth spreads, individuals (i) depend less on others to access critical resources (since they already own them) and (ii) find it beneficial to specialize and, in doing so, erect communication barriers. This explains why entrepreneurship resurges in fully developed countries: it becomes optimal to substitute specific knowledge for talent.

One can make an argument that increases (and respectively decreases) in the number of firms are self-reinforcing. This idea is implicit in Grossman and Helpman’s (2002) work on the trade-off

between coordination and specialization in vertical integration. They predict outsourcing when there is a highly competitive final-goods market, and cost savings from specialization within the supply chain are substantial (large enough to outweigh hold-up problems that arise from trading with specialized suppliers). On the other hand, vertical integration is favored when greater coordination creates value (final output price is very sensitive to the specifications of intermediate goods) and avoids bargaining issues (unique, difficult-to-substitute intermediate goods), or simply reduces fixed costs. Thus, when vertical separation occurs, it intensifies competition both downstream and upstream, and outsourcing (i.e. further separation) becomes more likely. Vertical integration reduces competition and thereby the costs of integration.

This mechanism may be another reason for the robustness of the recent trend toward more entrepreneurship in developed economies. The primary forces are, in our view, changes in the wealth distribution and the supply of infrastructure, which lead to changes in specialization and adaptation choices and ultimately changes in the number of entrepreneurs. Once this process is underway, the effect of increasing competition could strengthen it further.

The emergence of large commercial enterprises that employ and manage thousands of people is historically a young phenomenon that coincides with the dawn of the industrial age. Previously, small production units and a high share of self-employment dominated the commercial sector throughout much of human history (state bureaucracies, militaries and religious organizations were exceptions).

Various developments in England that preceded the industrial revolution removed many factors that had made collaboration costly. These changes eventually enabled low-skilled laborers to access the resources of the most talented and best-endowed entrepreneurs by joining their firms as wage workers. The English shared a common language and culture that was facilitated by a common religion and increasing levels of literacy and numeracy (Clark 2007, pp. 178-81). This circumstance enabled them to communicate more effectively and reduced management costs. In addition, urbanization saw large numbers of peasants migrate into the developing cities. The share of the English population living in cities increased from less than 5% in 1300 to more than 20% in 1800 (de Vries 1984, p. 39). This was crucial for mass employment in the newly emerging industrial firms to become feasible, and it provides a vivid example of the deliberate

choice to homogenize: people consciously chose to move to the cities, thereby drastically reducing the management costs of entrepreneurs who offered them jobs. Migration often paid off: there was a persistent wage premium for living in the maritime cities of northwestern Europe compared to the countryside (Allen 2001, 2003; Broadberry and Gupta 2006).

At the same time, the wealth base had gradually increased over the preceding ages and made capital and technologies available to more individuals. This is reflected, for example, in the decreasing returns to land and in the fact that rent charges dropped from around 10 percent in the 12th century to modern levels of around 4-5 percent by the eve of the industrial revolution (Clark 2007, pp. 168-170).

Some of those who now had access to capital turned out to be very able entrepreneurs. These individuals began to specialize in new, capital-intensive technologies and tapped into the growing middle class markets that emerged for a variety of products and luxury goods, ranging from clothes to books to crockery to watches etc. (Shammas 1990, McKendrick et al 1982, Berg and Clifford 1999, Berg 2005). Both developments together removed many of the obstacles to the formation of larger organizations and eventually enabled low-skilled, relatively homogeneous laborers to access the resources of the most talented and capitalized entrepreneurs by joining their firms as wage workers. As a result, average firm size increased, the share of relatively unproductive entrepreneurs in the population shrank, and large increases in income became feasible.

The exponential increase in incomes following the industrial revolution led to the satisfaction of basic needs among consumers in post-industrialized economies and to the emergence of products, services and small niche markets that targeted ever more individualistic consumer preferences. This development was accompanied by a continuing flow of innovations and growth of the knowledge stock, which finally made it possible for a large number of people to profit from specialization. They now had both the required wealth level to own and implement new production technologies, as well as new markets of consumers with sufficient willingness to pay for niche products.

The new opportunities strengthen the incentive to specialize and to acquire the knowledge and resources needed to exploit them. These investments in specialization in turn raise management

costs and limit the scope for collaboration in larger organizations. Although infrastructures are now highly developed, specialization and the greater heterogeneity it implies reduce the advantages of large firms and induce a tendency toward small-scale entrepreneurship. This explains the trend toward smaller firms and more entrepreneurship in highly developed economies.

5. Is There Dynamic Content in Transaction-Cost Economics?

Chandler (1992) wrote: "Just as I find the earlier growth of the industrial firm difficult to explain fully in terms of transactions, agency and other information costs, so I find it hard to explain the recent process of expansion and contraction with these same concepts. The answer requires a consideration of how the enterprises evolve their organizational capabilities ..." The theory we presented above is a step in this direction. The problem with applying the Coase-Williamson transaction-cost theory of the firm dynamically is that events which lower transaction costs usually lower them both inside and outside the firm. Therefore, they do not clearly predict changes in the boundary of the firm.

Coase (1937) originated the study of firm boundaries by providing a definition of what constitutes a firm and how size and organization of firms are determined. He viewed the firm as a contract that gives the buyer of a production factor control rights to direct the use of that factor: "The details of what the supplier is expected to do is not stated in the contract but is decided later by the purchaser. When the direction of resources ... becomes dependent on the buyer in this way, that relationship which I term a "firm" may be obtained." (Coase 1937: p. 392). Coase found this definition suitable because it represents the "essence of the legal concept of employer and employee" (p. 404). The unit of analysis is a transaction, and the boundary of the firm is determined by the relationship between market transaction costs and internal transaction costs (i.e. management costs). If the former exceed the latter, integration is preferred to outsourcing through the market.

Market transaction costs include the effort to discover what prices are, the costs of negotiating and concluding a contract, and paying taxes that apply to market transactions but not to

transactions inside a firm (e.g. sales tax). One might also include potential losses from giving up economies of scale, e.g. in the purchase or rental of capital inputs. Management costs, on the other hand, include any efficiency loss due to spatial distance between agents, costs related to principal-agent problems (i.e. shirking, monitoring costs, costs of negotiating and enforcing employment contracts) and bureaucracy. Hence, management costs limit collaboration within an organization, and transaction costs limit the extent to which spot markets can be used to trade productive resources. Both concepts together make the concept of a firm accessible to “the idea of substitution at the margin” (Coase 1937, p. 386).

This is a possible starting point for discussing the optimal number of entrepreneurs. If the firm is the group of individuals who cede authority to the entrepreneur under an incomplete contract, more entrepreneurs in a given economy imply, on average, more but smaller firms.

Various external factors can influence transaction and management costs. Improvements in infrastructure, such as building a road between two towns, will tend to decrease both types of cost. Improvements in the legal system that help to enforce property rights and contracts will likely do the same. The invention of book printing, the telephone, the internet, and improvements in general education conceivably had similar effects. Without belaboring the point, it is unclear in which direction these changes move the boundaries of firms and transform the efficient organization of workers in an economy.

Recent influential work has not so much emphasized the coordination problems stressed by Coase, but focused instead on the role of firm boundaries in providing incentives, in particular with regard to the “hold-up problem”. The hold-up problem refers to a situation where at least one party has to make a relationship-specific investment to enter into negotiations with another (e.g. drive to the factory), which is sunk in bargaining (the factory owner can still pay the laborer only the opportunity cost of her time, turning the overall transaction into a loss for her). Integration (factory hires the laborer on a contract) can resolve the hold-up problem.³ Property

³ The example here already suggests one limitation of transaction cost theory to explain the boundaries of the firm in our sense. Integration alone would not enable the relationship-specific investment between the worker and the firm in this example because the traveling costs of the worker would remain positive. If the productivity of the worker would not exceed her opportunity costs and the travel expenses, she could not profitably work for the firm,

rights theory considers how the initial assignment of asset ownership affects the efficiency of outcomes.

Williamson (1979, 1985) argued that the critical dimensions that determine the cost of transacting through the market vs. internally are (1) uncertainty, (2) the frequency with which transactions recur, and (3) asset specificity, the degree to which durable, transaction-specific investments are required to enable supply at the least cost. The cost advantage of internal organization relative to market transactions, and therefore the tendency toward larger firms and fewer entrepreneurs, is conjectured to increase in these aspects.

Which historical factors would lead to the observed U-shape between entrepreneurial activity and economic development? Technological progress, the increase in literacy and education, consumers' growing purchasing power and improvements in physical, legal and institutional infrastructure all typically accompany economic development. But their implications for uncertainty, frequency and asset specificity of transactions are unclear or conflicting.

For example, a better legal infrastructure reduces the uncertainty of transaction via the market, but also the uncertainty of organization-internal transactions that relates to the principal-agent problem. Inventions such as that of the internet decrease uncertainty by making information available at a much lower cost, but better search capabilities also facilitate an increase in the frequency of transactions. Technological progress could lead to more idiosyncratic job tasks (e.g. the need for an IT manager in a firm) or to a focus on a few core functions (e.g. outsourcing of accounting to a specialized provider via the internet). The net effect of economic development on the efficient boundaries of firms is ambiguous if we rely only on transaction cost arguments. There is no simple explanation in these terms for the historical trend in firm sizes and business ownership rates.

A transaction-costs argument that might have dynamic content was put forward by Tadelis (2002). His focus is on a trade-off between ex-ante and ex-post procurement costs: ex-ante cost is the contracted price for a specific design; ex-post cost is the expense of adapting the product

independent from who pays the travel expenses. The only way to enable cooperation in this case would be if the worker relocates closer to the firm, thereby reducing spatial distance and the geographic heterogeneity between the employer and the employee.

afterward to the buyer's needs. Market transactions allow for more competitive prices, hence lower ex-ante cost, at the expense of higher ex-post cost. More complex products (that would require more adaptation), are therefore produced in-house, while less complex products are procured through the market. If technological advances shift the product mix toward more complexity, integration and a smaller number of firms is implied. But in order to fit this argument to the U-shape, one would need a cycle of first decreasing, then increasing, and then again decreasing product complexity, for which there does not seem to be any evidence.

Property rights theory can be viewed as an extension of transaction cost economics that has tried to formalize some of the implications of the former, but it employs a different definition of the firm (Grossman and Hart 1986, Hart and Moore 1990, Holmström and Tirole 1991, Holmström and Milgrom 1991, 1994, de Meza and Lockwood 1998). In the property rights literature, the firm is defined as a set of non-human assets under common ownership. If two different assets have the same owner, they are understood to belong to a single, integrated firm; if they have different owners, then there are two firms, and transactions between them are carried out via the market (Holmström and Roberts 1998). The ownership pattern that maximizes aggregate surplus then defines the scope of the firm. Consequently, it will be efficient to have larger firms in cases where joint ownership of assets maximizes surplus. The market is the outside option for asset owners when they exit collaboration.

The concern of property rights theory is the effect of ownership of non-human assets on incentives to make relationship-specific investments. The theory is static in nature; hence it is not clear whether and how an optimal ownership structure would actually emerge. Results in the property rights literature are also very sensitive to assumptions about the nature of the relationship between parties (Holmström and Tirole 1991, Holmström and Milgrom 1991, 1994) and whether cooperation between parties is conceived as an *inside* option or an *outside* option to spot-market trading (de Meza and Lockwood 1998). This seems difficult to reconcile with the stable, long-term trends in entrepreneurial activity over the course of economic development. Most importantly, ownership rights for non-human assets are not equivalent to management authority and do not have clear-cut implications for the formal employment relationships that are central to our problem. One ownership structure might correspond to different arrangements of management authority among partners. Conversely, different ownership structures are possible

with the same arrangement of management authority. Thus, the theory has little to say about the optimal share of entrepreneurs versus wage earners in an economy.

Lafontaine and Slade (2007), in their review of the empirical evidence for transaction-cost and property-rights theories of vertical integration, conclude that predictions from transaction cost economics are supported in backward integration, whereas predictions from property-rights theory perform well in forward integration. We do not reject the significance of these theories for explaining differences in the firm size distribution between industries at any given time. However, we believe that one must look to other factors in order to understand long-term trends.

There seems to be little potential to infer firm size dynamics from any transaction-based explanation of firm boundaries, since it is difficult to determine the effect particular developments have on firm-internal vs. -external transaction costs. The property rights approach cannot really comment on empirical observations about the allocation of individuals to wage- vs. self-employment, since its unit of analysis is a non-human asset. A dynamic theory of the firm as a human organization needs to place individuals and their choices at the center.

6. Is More Entrepreneurship Always A Good Thing?

Given differences in talent and positive management costs, there will be, at every point in time, an optimal number of entrepreneurs. Differences in talent provide scope for complementarities, which in turn provide the rationale for setting up organizations and management structures that maximize joint output. Modern economies have come a long way in reducing management costs through infrastructure improvements. But management costs have not entirely vanished, if for no other reason than that cognitive limits constrain communication with those of different professional backgrounds. They limit the extent to which individuals can exploit complementarities and work together in one organization.

The trade-off between the benefits of complementarities among individuals and the costs of exploiting these benefits determines the optimal size and structure of organizations, and the optimal share of entrepreneurs. If this share were exceeded, society would not sufficiently take advantage of its most talented, who can offer workers productive and well-paid tasks. In the

absence of a theory that explains changes in this efficient number, the observed U-shape of entrepreneurial activity would suggest a degree of inefficiency in some countries and periods of time. We argue that this is not necessarily the case. The U-shape can be rationalized as an efficient process that responds to changes in infrastructure and wealth through people's specialization and adaptation choices.

While we view population homogeneity and heterogeneity as being affected by individual choices, there are constraints to the adjustments that can happen through this mechanism: a lack of basic transportation, communication and legal infrastructure will tend to require very small production units, since management costs (due to limited access and contract enforceability) are prohibitively high.

These types of constraints are empirically identifiable and relate to what we already know are factors inhibiting growth: a lack or misdirection of public infrastructure investments, dysfunctional rule of law, caste-like social stratification, etc. Improvements along these lines will release the binding restriction and, in principle, enable larger firms. In developed and threshold economies, where these hurdles have generally been overcome, the efficiency of the firm size distribution then depends mainly on incentives to adapt or specialize. To some extent, they are affected by policy, e.g. the tax code, which has a direct impact on returns to firm ownership and an indirect impact on product demand through the wealth distribution.

Many governments attempt to encourage entrepreneurial entry because business creation is viewed as something akin to a public good, since it generates new jobs. However, there is no reason to believe that such initiatives will necessarily lead to more productive, i.e. better-paid, jobs. There are already profit incentives to search for the best uses of labor; tax advantages for business owners (or discounted finance for entrepreneurship) instead may cause overentry by individuals who would contribute more as workers than as entrepreneurs. Similarly, distorted social security systems that give preferential access to employees and disadvantage entrepreneurs create the opposite bias: some "natural" entrepreneurs seek safer wage work instead, in order to take advantage of the associated benefits. In addition, employers will tend to outsource labor by forcing employees into self-employment.

Even when there are impediments to entrepreneurship that governments seek to correct through business-friendly policies, the focus should probably be on the primary causes (insufficient infrastructure and weak institutions), rather than on directly manipulating the business ownership rate. It is easy to lose sight of trends that work toward a desirable equilibrium number of entrepreneurs and then have policies in place that distort the size distribution of firms.

7. Conclusion

Implicit in many policy debates and programs is the view that entrepreneurship needs to be encouraged. Empirically, the share of entrepreneurs in the population tends to decline throughout economic development, significantly so since the industrial revolution, and only increase in advanced economies. This phenomenon is so well established that it is commonly referred to as the U-shape of entrepreneurial activity. Does the U-shape reflect temporary inefficiencies that policy must fix, or is the optimal boundary of the firm, and therefore the efficient number of entrepreneurs, evolving?

It is difficult to explain the U-shape using the traditional theory of the firm. Technological and social innovations on the eve of the industrial revolution in England, and globally ever since, reduced firm-external transaction costs and might seem to favor a steady increase in the number of entrepreneurs. But these changes would also be expected to reduce management costs within firms, and we have no guidance on which effect is likely to dominate. From a Coasian perspective, it is not obvious why firms started to grow in this period of time. An important factor is that the knowledge and the capital that were required in order to operate the available technologies efficiently were not available to most people. They were specific and attached to certain individuals, and the only way to access these valuable resources was by joining a firm that was operated by an advantaged entrepreneur. As a result, larger firms replaced smaller, craft-based ones (e.g. weavers), which previously produced similar products.

One can explain dynamics in the number of firms by recognizing that the size of firms depends on individual choices to specialize or adapt, with an early incentive to adapt (by acquiring uniform skills, such as a common language and basic literacy, and by co-locating in cities) and a

late-stage incentive to specialize (through advanced education). Adaptation lowers management costs and produces larger firms that take advantage of the resources and ideas of a few well-placed entrepreneurs. Specialization raises management costs and leads to smaller firms that are better able to exploit advanced technologies and niche markets. We argue that the timing of adaptation and specialization incentives depends on wealth patterns that evolve predictably.

Distortions occur when the development of critical infrastructure does not keep pace: for example, when property rights are absent or not upheld, and social discrimination restricts access to occupations. It is primarily these legal and social barriers that policy, especially in developing countries, has to address. On the other hand, policy in developed economies that is aimed at increasing the rate of entrepreneurship (e.g. tax advantages, discounted financing) runs the risk of encouraging an inefficiently large number of small firms, in which too many decisions are made by relatively untalented leaders.

References

- Acemoglu, Daron (2002), Technical change, inequality, and the labor market, *Journal of Economic Literature* **40** (1), 7-72.
- Acemoglu, Daron (1998), Why do new technologies complement skills? Directed technical change and wage inequality, *Quarterly Journal of Economics* **113** (4), 1055-89.
- Acs, Z.J., D.B. Audretsch and D.S. Evans (1994), Why does the self-employment rate vary across countries and over time? Discussion Paper No. 871, London: CEPR.
- Allen, Robert (2001), The great divergence in European wages and prices from the middle ages to the First World War, *Explorations in Economic History* **38** (4), 411-447.
- Allen, Robert (2003), Progress and poverty in early modern Europe, *Economic History Review* **56** (3), 403-443.
- Berg, Maxine (2005), *Luxury and Pleasure in Eighteenth-Century Britain*. Oxford and New York: Oxford University Press.
- Berg, Maxine and Clifford, Helen (1999), *Consumers and Luxury: Consumer Culture in Europe 1650-1850*. Manchester and New York: Manchester University Press.
- Broadberry, Stephen and Gupta, Bishnupriya (2006), The early modern great divergence: Wages, prices and economic development in Europe and Asia, 1500-1800, *Economic History Review* **59** (1), 2-31.
- Bosma, N., K. Jones, E. Autio and J. Levie (2008), Global Entrepreneurship Monitor, 2007 Executive Report, Babson College, Babson Park, MA, US / London Business School, London, UK.
- Carree, M.A., A.J. van Stel, A.R. Thurik and A.R.M. Wennekers (2002), Economic development and business ownership: an analysis using data of 23 OECD countries in the period 1976-1996, *Small Business Economics* **19** (3), 271-290.
- Carree, M.A., A.J. van Stel, A.R. Thurik and A.R.M. Wennekers (2007), The relationship between economic development and business ownership revisited, *Entrepreneurship and Regional Development* **19** (3), 281-291.
- Chandler, A. D. (1992), Organizational capabilities and the economic history of the industrial enterprise. *Journal of Economic Perspectives* **6**, 79-100.
- Clark, Colin (1960), *The Conditions of Economic Progress*, Third Edition, London: MacMillan & Co.

- Clark, Gregory (2007), *A Farewell to Alms – A Brief Economic History of the World*. Princeton and Oxford: Princeton University Press.
- Coase, Ronald H. (1937), The nature of the firm, *Economica* **4** (16), 386-405.
- De Meza, David and Lockwood, Ben (1998), Does asset ownership always motivate managers? Outside options and the property rights theory of the firm, *Quarterly Journal of Economics* **113**, 361-386.
- De Vries, Jan (1984), *European Urbanization, 1500-1800*. London: Menthuen.
- Edvinsson, Rodney (2005), *Growth, Accumulation, Crisis; With New Macroeconomic Data for Sweden 1800-2000*, Department of Economic History, Stockholm University, Stockholm: Almqvist & Wicksell.
- Fabre, Virginie, Condemi, Silvana, and Degioanni, Anna (2009), Genetic evidence of geographical groups among Neanderthals, *PLoS One* **4**, e5151, doi:10.1371/journal.pone.0005151.
- Fain, T.S., 1980, Self-employed Americans: their number has increased, *Monthly Labor Review* **103**, 3-8.
- Finlayson, Clive and Carrión, José (2007), Rapid ecological turnover and its impact on Neanderthal and other human populations, *Trends in Ecology and Evolution* **22**, 213-222.
- Finley, M.I. (1973), *The Ancient Economy*, London: Chatto & Windus.
- Freytag, A. and A.R. Thurik (2007), Entrepreneurship and its determinants in a cross-country setting, *Journal of Evolutionary Economics* **17**, 117–131.
- Grossman, G. M. and Helpman, E. (2002), Integration versus outsourcing in industry equilibrium. *Quarterly Journal of Economics* **117**, 85-120.
- Grossman, Sanford and Hart, Oliver (1986), The costs and benefits of ownership: A theory of lateral and vertical integration, *Journal of Political Economy* **94**, 691-719.
- Hart, Oliver and Moore, John (1990), Property rights and the nature of the firm, *Journal of Political Economy* **98**, 1119-1158.
- Hart, O. and Moore, J. (2005), On the design of hierarchies: coordination versus specialization. *Journal of Political Economy* **113**, 675-702.
- Holmström, Bengt and Milgrom, Paul (1991), Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design, *Journal of Law, Economics, and Organization* **7**, 24-51
- Holmström, Bengt and Milgrom, Paul (1994), The firm as an incentive system, *American Economic Review* **84**, 972-991.

- Holmström, Bengt and Roberts, John (1998), The boundaries of the firm revisited, *Journal of Economic Perspectives* **12** (4), 73-94.
- Holmström, Bengt and Tirole, Jean (1991), Transfer pricing and organizational form, *Journal of Law, Economics, and Organization* **7**, 201-228.
- Kuhn, Steven and Stiner, Mary (2006), What's a mother to do? The division of labor among Neandertals and modern humans in Eurasia, *Current Anthropology* **47**, 953-980.
- Kuznets, Simon (1971), *Economic Growth of Nations, Total Output and Production Structure*, Cambridge, MA: Harvard University Press / Belknap Press.
- Lafontaine, F. and Slade, M. (2007) Vertical integration and firm boundaries: the evidence. *Journal of Economic Literature* **45**, 629-685.
- Lazear, E. P. (2004), Balanced skills and entrepreneurship. *American Economic Review* **94**, 208-211.
- Lazear, E. P. (2005). Entrepreneurship. *Journal of Labor Economics* **23**, 649-680.
- Loutfi, M.F. (1992), An overview of self-employment in Europe; nature, trends and policy issues, in: P. Leighton and A. Felstead (eds.), *The New Entrepreneurs; Self-Employment and Small Business in Europe*, 41-68, London: Kogan Page.
- McKedrick, Neil, Brewer, John and Plumb, J.H. (1982), *The Birth of a Consumer Society: The Commercialization of Eighteenth-Century England*. London: Europa.
- Phillips, J.D. (1962), The self-employed in the United States, *University of Illinois Bulletin* **88**, Bureau of Economic and Business Research, Urbana: University of Illinois.
- Roessler, Christian and Koellinger, Philipp (2009), Entrepreneurship, firm formation and organizational design, *Tinbergen Institute Discussion Paper* 2009-003/3.
- Shammas, Carole (1990), *The Pre-industrial consumer in England and America*. Oxford et al.: Oxford University Press.
- Steinmetz, George and Erik Olin Wright (1989), The fall and rise of the petty bourgeoisie; changing patterns of self-employment in the postwar United States, *American Journal of Sociology* **94** (5), 973-1018.
- Stel, André van and Martin Carree (2004), Business ownership and sectoral growth; an empirical analysis of 21 OECD countries, *International Small Business Journal* **22** (4), 389-419.
- Storey, D.J. (1994) *Understanding the Small Business Sector*. London/New York: Routledge.
- Tadelis, S. (2002), Complexity, flexibility, and the make-or-buy decision. *American Economic Review*, **92**, 433-437.

- Wennekers, S. (2006), *Entrepreneurship at Country Level; Economic and Non-Economic Determinants*, Rotterdam: ERIM,
<http://publishing.eur.nl/ir/repub/asset/7982/EPS2006081ORG9058921158Wennekers.pdf>.
- Wennekers, Sander, Andre van Stel, Martin Carree, and Roy Thurik (2010), The relationship between entrepreneurship and economic development: is it U-shaped?, *Foundations and Trends in Entrepreneurship*, forthcoming.
- Williamson, Oliver (1979), Transaction-cost economics: The governance of contractual relations, *Journal of Law and Economics* **22**, 233-261
- Williamson, Oliver (1985), *The Economics Institutions of Capitalism*, New York: The Free Press.
- Wong, Kate (2009), Twilight of the Neandertals, *Scientific American* August 2009, 34-39.
- Yamada, G. (1996), Urban informal employment and self-employment in developing countries: theory and evidence, *Economic Development and Cultural Change* **44**, 289-314.