Entrepreneurship and the economic theory of markets

Article

Accepted Version

Entrepreneurship and the economic theory of markets

ABSTRACT

Entrepreneurship studies has an uneasy relationship with mainstream economics. This paper attempts to bridge the gap by presenting a formal mathematical theory of entrepreneurship, focusing on competing entrepreneurs operating in a segmented market. The entrepreneur is modelled as an innovative market-maker, discovering opportunities for trade. Each market segment is characterised by a maximum price that customers are willing to pay, and a quantity that they are willing to buy at that maximum price. The model can solve for the outcome of complicated market structures using a simple technique, namely linear programming.

A remarkable feature of the model is that it re-ignites historic debates in mainstream economic theory that have never been fully resolved, and are largely ignored in contemporary economic literature. It shows that markets normally have two prices - not one – namely a wholesale price and a retail price - and that, owing to segmentation, prices are not normally uniform but dispersed. The mathematical model is sufficiently simple that it does not require calculus, and the method of solution is so straightforward that the requisite techniques can be freely accessed on a laptop. [183 words]

KEYWORDS

ENTREPRENEUR, THEORY, MARKET, TRADE, SEGMENTATION, PRICE, PROFIT

Introduction

Entrepreneurship studies embraces a wide field of research, and within this field the formation and growth of small firms, and the management challenges it creates, occupy much of the literature [Brueghel, Andersson and Eklund, 2022]. By contrast the economic theory of entrepreneurship has not attracted as much attention as it deserves [Hebert and Link, 1982]. This paper attempts to rebalance this situation, by presenting a novel theory that synthesises the main insights of existing theories of entrepreneurship. The paper is addressed to both entrepreneurship scholars and to business and industrial economists.

The paper focuses on the role of entrepreneurs in a market economy. It presents a theory of innovation, imitation and competitive rivalry in a segmented global market. Within this context entrepreneurs are the key agents of change. The paper is unusual in setting out a formal mathematical model of entrepreneurship, of a type that previous entrepreneurship scholars have largely neglected. A formal model is essential if this paper is to prove its point in a way that convinces economists as well as entrepreneurship scholars [Baumol, 1968].

There are three advantages of a formal model.

- Key information about a market, namely the quantities traded, the prices, and the profits, are all measurable quantities which are best explained in terms of other measurable factors such as the intensity of demand and the cost of production.
- Mathematics provides a rigorous logic that can be applied to the analysis of quantities. It does not suffer from the ambiguities of language, where words often carry connotations that are value-laden, and whose meaning is defined according to the context in which they are used.
- Mainstream economists make extensive use of models, and to convince these economists of the value of an alternative perspective on markets it is useful to present a new model that has the same standard of logical rigour as the models they at present use.

The paper is in six parts. Parts One, Four and Six are essential reading. Part One introduces the topic. Part Two presents a formal model which is the analytical core of the paper. The model involves algebraic equations and inequalities, and is solved using linear programming. This part establishes the logical rigour of the theory, but could be omitted on a first reading. Part Three presents some case studies that illustrate the wide-ranging scope of the model. Readers who have not read Part Two may wish to skim-read Part Three in order to see how the model is applied. Part Four reflects on how the model compares with alternative approaches to the economy, from the formative period of economic theory down to the present day. Part Five examines the social fabric within which entrepreneurship is conducted, with an emphasis on institutions such as clubs and cartels. It shows that co-operation as well as competition is a vital element in entrepreneurial activity. Conclusions and policy implications are summarised in Part Six, with a special emphasis on the teaching of entrepreneurship and its relationship to economic theory.

PART ONE: OVERVIEW

History of thought

Three main schools of thought have addressed the role of the entrepreneur in a market economy:

- The Austrian school, which focuses on speculation, arbitrage and the market process [Foss and Klein, 2002, 2012, 2020; Harper, 1996; Kirzner, 1973, 1979; Klein, 2010; Ramoglou, 2021; Ramoglou and Tsang, 2018];
- The Schumpeterian school, which focuses on the dynamics of innovation and growth [Arthur, 1989; Backhaus, 2003; Baumol, 1990; De Vecchi and Stone, 1995; Langlois and Robertson, 1995; Schumpeter, 1934, 1939; Swedberg, 1991]; and
- Mainstream economics, which is preoccupied with the determination of equilibrium price and equilibrium output. In mainstream models entrepreneurs are often identified with the owners of firms, and their market-making role is ignored altogether [Solow, 1956].

While the Schumpeterian entrepreneur innovates a novel product or technology, the Austrian entrepreneur 'fills a gap' in existing markets. The theory presented in this paper synthesises these two approaches by developing a new approach in which they can both be embedded. It involves a formal model that has many parameters. Because the theory is so general, and the model has so many parameters, applications can be tailored to the many and varied forms of entrepreneurial activity. The theory is an economic theory and uses basic concepts from mainstream theory, such as scarcity, opportunity cost and Pareto efficiency.

The theory focuses on the role of the entrepreneur in an economy with segmented markets, as indicated above. Segmentation is a specific type of market imperfection which has received less attention than it deserves. Segmentation partitions a market into individual segments with limited communication and limited trade between them. Segmentation can be analysed at a global, national, regional or local level. It may be created by geographical distance, linguistic differences, cultural differences and various other factors, as explained below. Segmentation can create many segments or just a few. It can also be tiered, as when national segments are composed of regional segments, and regional segments of local segments, and so on. There can also be different dimensions of segmentation that interact, as when each specific region is composed of different occupational groups that do not communicate fully with each other [Bourdieu, 1977].

Segmentation introduces a vital element of realism into the economic models described above. Different competitive conditions may prevail in each segment. Some entrepreneurs may be involved in several segments, each of which may have its own distinctive properties. When considered as a whole a segmented market contains multiple entrepreneurs, each focused on a different set of one or more specific segments. Some segments may be monopolised, but others may be competitive because they are served by several entrepreneurs, i.e. the segments entered by different entrepreneurs overlap. The monopolistic segments may generate substantial profits, but the competitive segments may generate little or no profit. When the gains from trade are not appropriated by an entrepreneur they are appropriated by customers or producers instead.

Outline of the model

The rationale of the model is easy to explain in general terms. It disposes of the fictional Walrasian auctioneer, who is so prominent in mainstream economic models of the market [Walras, 1954]. It replaces them with a set of competing entrepreneurs. Each of these entrepreneurs connects up specific segments of demand with specific segments of supply. They act as intermediators, buying from producers and reselling to their customers. They are

motivated by profit, which is appropriated by a margin that they set between their buying price and their selling price. This margin covers their operating costs as well.

When markets are segmented information collected about one segment does not necessarily provide information about other segments too. Connecting up a segment of demand with a segment of supply provides an opportunity for profit. Entrepreneurs are people who have chosen to specialise in seeking out opportunities of this kind. Different entrepreneurs may have access to different information; in particular, they may have access to information about different segments of the market. If these segments overlap then there is competition between them and if they do not then they have a monopoly.

The initial condition is a no-trade situation in which none of the opportunities for trade have been recognised. The most alert entrepreneur enters first, but typically fails to exploit all of the available opportunities. Their information is incomplete and therefore they enter only specific segments with which they are familiar. Their entry catches the attention of other entrepreneurs, however, who follow their example. In some cases the followers compete directly with the entrant, and in other cases they enter market segments that the first-mover has overlooked. As further entry occurs, followers begin to compete with earlier followers, competition intensifies, and some entrepreneurs may be driven out of the segments that they entered to begin with.

Where two or more entrepreneurs serve the same segment of demand or procure from the same segment of supply they compete with each other; otherwise, as the only supplier to some specific customer segment, they have monopoly power or, as the only purchaser from some specific producer, they have monopsony power. Under monopoly an entrepreneur can, under certain conditions, appropriate all the gains from trade as profit, but, as indicated above, under competition they cannot; some or all of the gains from trade will accrue to customers or producers instead.

Without entrepreneurs there would be no markets, or at least no organised markets. Without some monopoly there would be no incentive for entrepreneurs to establish markets, but with too much monopoly producers and customers, i.e. the public in general, have little to gain. Society therefore faces a trade-off. If it offers monopoly power then it increases the supply of entrepreneurs, but if it encourages entrepreneurs to compete rather than to collude then entry may be deterred. This trade-off is widely recognised in the field of technological innovation, where it provides a rationale for patents with a limited life, but it is not so widely recognised in respect of ordinary markets.

'Market-making' entrepreneurs

The entrepreneurs who are the first to identify a trading opportunity may be described as 'market-making' entrepreneurs. A market-maker is essentially a first-mover in some segment of a market [Casson, 1982, 2005, 2010; Godley and Casson, 2016]. If different entrepreneurs focus initially on different segments then there may be several market-making entrepreneurs in the market as a whole. Whether they are competitors or monopolists depends on whether they expand into segments other than the one they initially entered, as indicated above.

Market-making entrepreneurs intermediate between buyers and sellers. They generate their profit by driving a wedge between the buying price and selling price in the market where they operate. This wedge – or 'mark-up'– is required to cover their operating costs. The margin

that remains after deduction of these costs represents their net profit. This highlights another feature of real-world markets that is missing from mainstream economics models – namely the divergence between retail price (at which product is sold to customers) and wholesale price (at which product is purchased from producers) or, equivalently, the mark-up (often expressed in percentage terms) applied by an entrepreneur to the wholesale price.

Market segmentation and barriers to trade

The assumption of market segmentation, which is key to the model, is a useful antidote to the assumption of perfect integration which is usually made in the economic analysis of markets, and especially in the analysis of national markets for a homogeneous product [Gould, 1980; Nooteboom, 2014].

Perfect integration requires both perfect knowledge of trading opportunities and the absence of any barriers to trade. The crucial role of entrepreneurs in collecting information in order to identify trading opportunities has already been noted [see also Hayek, 1945]. But even if entrepreneurs had perfect information there would still be barriers to trade to be overcome. An entrepreneur can overcome cognitive obstacles caused by a failure to recognise potential sources of demand and supply, but they cannot overcome the basic barriers that impede the conduct of trade. Recognition of these barriers is a key element in any account of the role of entrepreneurs in generating trade. Indeed, the ability to resolve the problems created by barriers to trade is one of the hallmarks of a successful entrepreneur.

It is useful to distinguish between 'hard' and 'soft' barriers to trade. Hard barriers include transport costs, tariffs, state regulations (e.g. quality controls) and the costs of currency exchange. Costs of communication by telegraph, telephone and internet are also potentially significant. Soft barriers relate mainly to the marketing of products to unfamiliar buyers. They include unanticipated differences in local customs, which may influence customer perceptions of the product and affect the way in which it is used; and unanticipated variation in the age and affluence of potential customers, which may affect the prices that they are willing to pay.

Differences in language create both hard and soft barriers. They create hard barriers in the negotiation of contracts, and soft barriers when advertising the virtues of some novel product or creating an image for some newly established brand.

There are three types of economic model in which barriers to trade are prominent, and these are itemised below.

International trade theory is a special type of spatial economic theory, and includes significant elements of regional economics and transport economics [Meade, 1952; Whitin, 1953]. It has always emphasised the importance of barriers to trade between national markets caused by transport costs, tariffs and non-tariff barriers such as quotas and national standards for product quality. In particular, the economic theory of tariffs represents an early exercise in the analysis of market segmentation. Another feature of international trade theory is its emphasis on the immobility of resources such agricultural land, mineral deposits, and the influence of local climate on economic activity. This is of great historical importance in analysing early colonisation and imperialism, and is also of considerable contemporary relevance in analysing the effects of climate change.

Labour markets also exhibit segmentation. This is mainly a consequence of workers' differences in aptitude and ability, as noted by Adam Smith [Meek and Skinner, 1973; Smith, 1776; West, 1969]. According to Smith, workers are most productive when they specialise, thereby creating different segments of the labour market, each based upon some specific set of skills. Workers with specific skills agglomerate at specific locations which afford access to the specific localised resources that their work requires. In this context intermediation of the labour market serves to match workers with specific skills to employers who offer suitable jobs.

Failure to match efficiently may cause 'frictional unemployment'. The role of the market-maker is played mainly by employment agencies, although social networks can be important too. Private agencies operate on a for-profit basis, but state agencies normally operate on a subsidised 'not-for-profit' basis. Private agencies may focus on more talented groups while state agencies focus on more vulnerable groups. Market-making is complicated by the fact that in civilised societies workers cannot be bought and sold; a common solution is for an agency to hire the worker and then re-sell their services to the ultimate employer. Other intermediating institutions include trades unions and professional associations, which bargain with employers on behalf of their members. Segmentation is also exacerbated by the geographical immobility of workers who may be embedded in an extended family or local community that supports them.

Product differentiation is a well-known form of market segmentation. Customers who favour one variety of a product may reject another variety, or only purchase it at a lower price. In other words, one variety of product is not a perfect substitute for another. In conventional economic models product differentiation is often measured by the positioning of a product along a line or circle, with the emphasis being placed on the 'distances' along the line between different varieties of product. Theories of product differentiation feature in industrial economics and the economic analysis of business strategy [Chamberlin, 1933; Hamilton and Richards, 2009; Hotelling, 1929]. The 'new trade theory' of the 1970s combined elements of product differentiation and barriers to trade to analyse trade flows in sophisticated modern products [Krugman, 1980]. Product differentiation is highly relevant to entrepreneurship studies, but it is not the specific focus of the present paper.

Because barriers to trade create segmentation they can lead to different prices in different segments of a market, as explained in more detail below. Seeking economic efficiency by 'getting the price right', as in the Chicago tradition of economics, is a difficult principle to apply when markets are segmented and different prices prevail in different segments [Stigler, 1957]. In practice prices are not uniform, as in mainstream economic theory, but dispersed.

Mainstream economics explains how consumer search can reduce price dispersion, but it does not explain how price dispersion arises in the first place. Because consumer search is costly it can only reduce price dispersion and cannot eliminate it altogether. Mainstream economics has therefore ignored price dispersion on somewhat spurious grounds. The main exception has been the theory of international trade, as noted above. However, even where price dispersion is indisputable, as in international trade, it is normally attributed entirely to barriers to trade and not at all to lack of information [Stigler, 1942, 1957, 1961]. The model presented in Part Two below shows that if both lack of information and barriers to trade are

taken as given then an efficient market outcome, conditional on these barriers, does not require a uniform price [Casson, 2024].

Market-makers and producers

To analyse the market-making process it is vital to distinguish between the entrepreneur as market-maker and the entrepreneur as producer. The market-maker is responsible for matching supply to demand. They typically identify a latent demand that other people have overlooked and then seek out a source of supply with which to satisfy this demand. In this way they 'make a market' for a product which did not previously exist within the segment that they serve. The producer is the source of supply; they typically employ labour and combine it with other resources (land, buildings, machinery, etc.) to generate their output. An entrepreneurial producer may decide to exploit a novel technology (see below) or simply recruit a team of workers to produce a product that market-makers will wish to buy.

The market-maker and the producer therefore fulfil different functions. The market-maker seeks out information on supply and demand and identifies an opportunity (as in the Austrian tradition) to connect them up. This involves advertising (and possibly displaying) the product to customers, setting the price, collecting payment and providing (where appropriate) after-sales service. They are therefore mainly concerned with the handling of information and managing the consignment of the product from the producer to the customer.

The producer, on the other hand, is mainly concerned with organising labour into teams and directing these teams to transform inputs, such as raw materials, into output of the product. The market-maker and the producer typically rely on different sources of information and require different skills. In order to incentivise themselves they normally take ownership of the product at the relevant stage. In this case the product will be produced by an independent producer and sold at 'arm's length' to an independent market-maker.

Integration

A market-maker may, however, be concerned about the quality of the product supplied, in which case they may prefer to integrate backwards and take ownership and control of the production process themselves. Conversely, a producer may be concerned that an independent market-maker might switch their orders to a rival producer, and so, to secure an outlet for their product, they may integrate forward and market their product themselves. This is called the 'make or buy', 'internalisation' or 'vertical integration' decision [Alchian and Demsetz, 1972; Buckley and Casson, 1976; Coase, 1937; Richardson, 1960; Williamson, 1975, 1985].

There are other well-known reasons for integration. An entrepreneur may have devised a product that embodies a new technology, or which is differentiated from potential rivals by a distinctive brand. Subcontracting production to an independent producer may run the risk of piracy, whereby the producer steals the technology or the brand and integrates forward into market-making themselves. In this case the entrepreneur may integrate backwards into production by acquiring ownership of the production unit they plan to employ.

Another possibility is that the producer owns an exhaustible resource which the entrepreneur expects to increase in price as it becomes depleted by rising demand for their novel product. They may also be concerned that rival entrepreneurs may enter the market and out-bid them

for supplies, or even acquire the resources for themselves. As a pre-emptive measure, they may therefore acquire ownership of the resource in order to monopolise access.

The question then arises as to whether the market-maker or the producer is the entrepreneur, or whether they are both entrepreneurs. The answer to this question depends on the specific circumstances of each case. In terms of the model presented below, it is assumed that the market-maker, and not the producer, is the entrepreneur.

Firms and entrepreneurs

In mainstream neoclassical theory key decisions are taken by the 'firm', whereas in entrepreneurship theory they are taken by the entrepreneur. This paper follows the entrepreneurship tradition. Decisions are made by people, and the entrepreneur is a person and a firm is not. The firm is a legal institution that functions as a transactions hub. It hires labour, land and other resources, and sells product. It can also borrow money and issue shares to investors in return for interest and dividend payments. As a fictional person it can also outlive its founder, which may ameliorate issues of continuity and management succession [Commons, 1924; North, 1991; Hodgson, 2015].

The management of the firm is a key responsibility of the entrepreneur. The raising of capital to finance the firm is also an entrepreneurial responsibility. Issues relating to the dilution of control, the structuring of debt (e.g. between equities and debentures) and the planning of management succession demand the attention of the entrepreneur because they are key to the survival and growth of the firm [Knight, 1921; Penrose, 1959]. These issues are not the focus of this paper, however; it is the relation of the firm to the market that is key.

The important point is simply that the integration of production and marketing, as described above, leads to a vertically-integrated firm. Similarly, an entrepreneur who serves multiple market segments from local plants may control a horizontally-integrated firm; indeed, if these plants are located in different countries then the firm will become a multinational. Likewise an entrepreneur that innovates multiple products may become the owner of a diversified firm. A highly successful entrepreneur could even become the owner of a diversified and vertically-integrated multinational firm [Baligh and Richartz, 1967; Buckley and Casson, 1976; Porter and Livesay, 1971].

Innovation

Innovations may be the outcome of scientific research, market research or spontaneous discovery, all of which are difficult to model in a mainstream economic context; furthermore, the appearance of a novel product requires the reformulation of individuals' utility functions, which creates additional problems for mainstream theory [Teece, 2012; Teece, Pisano and Shuan, 1997]. The theory presented below is designed specifically to accommodate innovation; it can therefore analyse change as well as statis. It also dispenses with utility functions, which makes it easier to accommodate innovation.

A typical example of innovation in a segmented global economy is as follows. An entrepreneur (a 'first-mover') trials an innovation in a specific country or region with which they are familiar, and plans to expand into other areas. Other entrepreneurs ('second-movers') witness their success, and contemplate entry, thereby becoming potential rivals. A second-mover may decide to imitate the first-mover by entering their home market to compete for

market share, while another may decide to launch a similar product in a foreign market with which they are familiar. In the first case competition is confined initially to a single segment of the overall market, while in the second case there is no competition because each entrepreneur serves a different segment of the overall market. The first-mover entrepreneur may subsequently decide to expand into the second-mover's segment, and the second mover may decide to expand into the first-mover's segment. As more entrepreneurs enter and the market globalises many different configurations can emerge, in which some segments remain monopolised through successful entry deterrence while others become highly competitive.

PART TWO: A FORMAL MODEL

A general model of entrepreneurship

The observations above lead directly to a general theory of entrepreneurship, based on a formal model. [For other 'general theories' see e.g. Baumol, 1994; Shane, 2003].

The model is a general equilibrium rather than a partial equilibrium model in the sense that it models the global economy rather than some subset of it (e.g. a region, or a nation state). The entrepreneur is portrayed as an agent of change, who either innovates a new product, launches an existing product in a new market segment, introduces the production of an existing product to a new location, or establishes a new linkage, or trade route, between an established production location and an established market segment.

The simplest version of the model focuses on a single entrepreneur who introduces a novel product to a single segment of the global economy. A more sophisticated model analyses the sequential entry of entrepreneurs who diffuse consumption and production to additional segments of the economy. Entrepreneurs compete against each other, both to discover new market segments and new sources of production, and to invade each other's segments in the expectation of either driving rivals out or sharing the market segment with them. The model can be extended to include multiple novel products, provided that there is limited interaction between them.

The basic principle is that entrepreneurs collectively and unintentionally maximise the global gains from trade subject to the constraints imposed by their ignorance of specific markets or specific locations of production. They do not normally appropriate all these gains for themselves, however. As explained above, where they compete for sales to the same market segment, some of the gains may accrue to their customers, and when they compete for supplies of the product some of the gains may accrue to the producers instead. This 'unintended outcome' of entrepreneurial rivalry is essentially an application of Adam Smith's 'invisible hand' to the organisation of global trade.

In very general terms the unintentional maximisation of the global gains from trade arises from a sequence of myopic increments in the gains from trade achieved through localised coordination between individual traders. When a potential buyer and a potential seller make contact through an intermediary all three potentially gain, because if they did not gain they would have no incentive to participate in the trade. At very least, one person will gain and the others will not lose (i.e. in the terminology of mainstream economics, Pareto-improvement will be achieved.)

The most likely person to make the gain is the intermediary, because they will normally have recognised the opportunity first, and set the terms on which the trades are made to favour themselves instead of the other parties. In the present context each intermediary is a specialist entrepreneur who typically undertakes many trades involving different people. They appropriate their gains by 'buying cheap and re-selling dear'. In a typical case an entrepreneur approaches a customer segment which they have correctly identified as having a need for a product which is not yet available, or which is already available but has not been offered to them by others. They also identify one or more potential producers of this product. They contract individually with the producer(s) and offer a general invitation to trade to potential customers, stipulating the price at which they are willing to supply which (given the accuracy of their information) is equal to the maximum price that the customers are willing to pay. This maximises the gains from trade subject to the information available at the time.

The gains from trade are measured in terms of a basic numeraire product for which everyone has an insatiable demand. However, the gains from trade generated by the model do not, strictly speaking, measure the increase in global welfare. There are three main reasons for this.

- Different people may derive different satisfaction (or utility) from the basic product. Because the price of a novel product is measured in units of the basic product, a person who values the basic product very little may value the novel product highly in terms of the basic product, even though they derive only limited satisfaction from it.
- The existence of any really 'basic product' for which there is insatiable demand is highly dubious. It could be said that 'money' is such a product, but money simply circulates, so that increasing the supply of money simply leads to price inflation rather than any increase in the real value of output. However, some sort of basic product is essential to the model, especially for accounting purposes.
- The aggregation of the gains from individual trades makes no allowance for inequalities in personal income, or for the sense of alienation and injustice that it can create. In view of the enormous profit that monopolistic entrepreneurs can potentially make, this is a serious practical limitation of interpreting the gains from trade as a measure of the increase in social welfare.

The model must therefore be judged solely by its explanatory power. This paper demonstrates that its explanatory power is potentially high. The model not only sheds new light on existing topics but also highlights other topics that have received less attention than they deserve in mainstream entrepreneurship literature.

Structure of the model

Consider an economy with a given set of entrepreneurs, each of whom is aware of one or more segments of demand for a given novel product. The entrepreneurs are a self-selected group; they have all committed to enter the market for this product. There may also be other entrepreneurs who were either unaware of the market or considered entering it but decided to stay out. These differences in entry decisions reflects the diversity of the personal judgments made by potential entrepreneurs. A comprehensive analysis of entrepreneurship would analyse these entry decisions [Casson, 2024], but that is not attempted here.

The economy is assumed to include just two products, namely a basic product and a novel product, which is the focus of the theory. Different segments of the market demand different amounts of the novel product and the customers in these different segments are willing to pay different prices for it. The only alternative to consuming the novel product is to consume the basic product, for which everyone has an insatiable demand. The basic product can be produced by anyone, and is produced under constant returns to scale. It is a unit of account and therefore always has a price of unity. People who do not produce the novel product, or act as entrepreneurs, all produce this basic product. This approach to economic modelling is quite well known, and is implicit in many economic models.

Each entrepreneur who has entered the market for the novel product is in touch with a given set of producers and a given set of customer groups. There is no communication between customer groups and no communication between producers, while all communication between customers and producers is intermediated by one or more entrepreneurs. Although in principle any entrepreneur could put any producer in contact with any customer group, in practice different entrepreneurs have different sets of contacts, although there may be considerable overlap between them. Each entrepreneur's personal set of contacts determines a set of potential linkages, or routes, by which each potential customer group can be served by the output of a potential producer through the intermediation of the entrepreneur. Each entrepreneur is a hub through which each of these routes passes. The solution of the model determines which linkages operate and how much product flows along each of them.

Each route may incur some access costs. Access costs are borne by entrepreneurs, and may arise either from the procurement of product from a producer or the supply of the product to a customer segment. Costs of procurement include the costs of negotiating with the producer, checking quality and taking delivery of the product. Costs of supplying customers are typically much greater because it is necessary to make contact with each individual customer within a given group. These costs include advertising, display, physical distribution, stock control, collecting payment and providing after-sales service. It is assumed for simplicity that access costs are directly proportional to the quantity traded along a specific route.

Some entrepreneurs may monopolise just a single customer segment and some may monopolise several. In addition, some segments may be served competitively. Competition may involve two or more entrepreneurs, depending on the number of entrepreneurs and the overall size of the market.

In general there will be N_1 customers, N_2 suppliers and N_3 entrepreneurs. This creates a very large number of possible market structures, $N = N_1 \times N_2 \times N_3$, each of which involves a different pattern of segmentation.

A simple example of market structure

A simple market structure is illustrated in Figure 1. In this case $N_1 = N_2 = N_3 = 2$, and so the total number of possible linkages is $N = 2 \times 2 \times 2 = 8$. People (individuals or groups) are represented by boxes and the arrows between them represent connections between producers and entrepreneurs and between entrepreneurs and their customers. The direction of each arrow represents the direction of the flow of the product. There are two customer groups, $A1$ and $A2$, neither of which is in touch with the other. There are two independent producers, $B1$ and $B2$, who are also not in contact with each other. There are two entrepreneurs, $E1$ and $E2$.

*E*1 is in touch with customer groups *A*1 and *A*2 and producers *B*1 and *B*2, while entrepreneur *E*2 is also in touch with both producers but only with a single customer group, *A*2. Thus only seven of the eight linkages operate in this particular case; the missing linkage is between *A*1 and *E*2. All the other applications of the model (including those presented below) are variants or extensions of this simple type of model.

FIGURE 1 HERE

It is assumed that each customer segment has a given total demand (known as the 'satiation demand') at a given uniform reservation price, and each potential producer has a constant unit cost up to a capacity limit. These are crucial simplifications, without which the method of solution will not work. They imply that both the demand schedule and supply schedule are rectangular. It must be recognised that all computable models of the economy place restrictions of some type on the form of demand and supply schedules. Marshall, for example, assumed constant elasticities, and this precedent is still followed in many neoclassical economic models today [Marshall, 1890]. Constant elasticity is a very strong assumption, however, and arguably no less restrictive, in certain cases, than the assumptions described above.

Note that the reservation prices and the quantities demanded may vary between the customer segments, and likewise the unit costs and capacities may vary between producers. Because of segmentation, therefore, there are far more parameters governing the overall levels of demand and supply than there are in a familiar Marshallian model.

Only entrepreneurs are aware of the opportunities for trade. No entrepreneur, once they have committed to their role, will ever pass up an opportunity for gain of which they are aware. In equilibrium, therefore, there is no unexploited opportunity to gain from trade along any recognised route. This means that the gains from trade are maximised, conditional on the information available to the entrepreneurs, i.e. they are maximised subject to the constraints imposed by the set of linkages available [Ingrao and Israel, 1990].

Method of solution

The maximisation of the gains from trade subject to constraints is a special case of an optimization problem. Because of the assumption of constant reservation prices and constant unit costs, all the constraints take a linear form. Likewise the total gains from trade are the sum of the individual gains from trade achieved by the various linkages, and therefore also have a linear form. The optimization problem is therefore a special type of linear programming problem.

Linear programming problems are easily solved, so that a solution to the market problem can always be found, although it may not always be unique. This solution determines which linkages operate in equilibrium, and how much of product is routed along each linkage. It therefore determines how much product in total is routed through each entrepreneur.

Linear programming models are typically solved using the simplex method. This method has a simple intuitive explanation involving an *N*-dimensional space. The demand, supply and entrepreneurial capacity constraints, when combined, define the frontiers of a convex set in a Euclidean hyperspace. This hyperspace has a number of dimensions equal to the number of possible routes linking any customer segment to any producer through any intermediating

entrepreneur. The maximum number of dimensions is $N = N_1 \times N_2 \times N_3$. The boundary of this convex set comprises flat sections. There are corners (or 'vertices') where several adjoining sections meet, while ridges connect adjacent corners to each other [Boulding, 1960; Dantzig, 1991; Dorfman, 1951; Dorfman, Samuelson and Solow, 1958; Koopmans, 1951]

Each linkage makes a specific contribution to the gains from trade. This contribution is equal to the customer reservation price in the destination market *less* the unit cost of production in the relevant production unit l*ess* any access costs incurred (measured as constant charges per unit traded). This contribution sets an upper limit on the amount of profit that can be appropriated from the linkage by the entrepreneur that operates it.

Each value of the maximand (i.e. the gains from trade) defines a plane in this hyperspace. Varying the value of the maximand generates a set of parallel planes, each of which corresponds to a specific total value of the gains from trade. The highest-value plane to touch the surface of the convex set determines the maximum attainable value of the gains from trade, while the location (coordinates) of the relevant corner (or ridge) determines the configuration of the trading system that supports it. A corner (or vertex) determines a unique solution, while a ridge determines a range of possibilities, all of which generate the same value of the gains from trade.

The value added by an individual entrepreneur can be derived by assessing the impact on the total gains from trade of constraining the linkages promoted by that entrepreneur to zero and recalculating the solution. The reduction in value indicates the maximum profit that can be appropriated by the entrepreneur, although this may exceed their actual profit, for reasons described below.

Prices and profit

In a conventional linear programming problem prices are calculated by solving a 'dual' to the problem set out above, but this will not determine the prices in the present case. In the present case the dual values determine the gains from trade that are generated by a marginal relaxation of constraints on the amount of trade that can be conducted along a given linkage. These gains may accrue to either the customers, the producer or the entrepreneur, depending on whose capacity constrains the overall capacity on a given route. In the version of the model presented above there was no capacity constraint on the ability of an entrepreneur to conduct trade, either as a whole or along any specific route. Thus no reward to entrepreneurs will be imputed by the dual problem; the dual will simply impute rewards to the customers or the producers, depending on whether there is excess supply or excess demand on a given route.

To allocate rewards it is necessary to introduce additional assumptions, which can be derived from basic economic principles. It is therefore assumed that, in the absence of access costs:

- If a customer group has excess demand for the product (i.e. their satiation demand exceeds the supply of product forthcoming at the reservation price) then the price remains at their reservation price however many competing suppliers there are.
- If a customer group has excess supply from a solitary entrepreneur at their reservation price then the price remains at the reservation price, notwithstanding the excess supply, because the entrepreneur concerned exercises monopoly power. Customers do not

exercise monopsony power because they are a fragmented group and do not have the same information about production cost as does the entrepreneur.

- If a customer group has excess supply from two or more entrepreneurs then the price is the break-even price of the marginal entrepreneur. Intra-marginal entrepreneurs price just below this to ensure that all their output is purchased by customers before they purchase from the marginal producer. As a result the marginal producer normally operates at below capacity (unless they supply another market too) while the intra-marginal producer operates at full capacity.

Another set of assumptions applies to producers. The thrust of these assumptions is somewhat different from the set above because each producer negotiates on behalf of their entire production team, while no customer group has a representative to negotiate on their behalf.

- If a producer has excess supply when selling to entrepreneurs then they charge their unit cost, or break-even price, however many buyers there are.
- If a producer has excess demand from a single entrepreneur then they still sell at their break-even price (i.e. unit cost) because the entrepreneur can exercise monopsony power.
- If a producer has excess demand from two or more entrepreneurs then they sell to the marginal buyer at the buyer's break-even price. The intra-marginal buyers (i.e. entrepreneurs who serve higher-price markets) offer just above this price in order to ensure that all their requirements are met.

These rules become somewhat more complicated where access costs are involved, but the same principles still apply. For further discussion see Casson [2024].

Example: *Competition for market share*

A simple example is illustrated in Figure 2. As before, there are two market segments to be served, *A*1 and *A*2. The customers are 'downstream' and the producers are 'upstream'. Both customers and producers are segmented by location. There is no direct contact between customers *A*1 and *A*2, or between producers *B*1 and *B*2, and no direct intermediation (i.e. arbitrage) between them either.

FIGURE 2 HERE

Customer segment *A*1 demands 60 units of product at a reservation price of 10, and segment *A*2 also demands 60 units of product but at a lower reservation price of 8. There are two potential producers, *B*1 and *B*2. *B*1 can produce 40 units at a unit cost of 4 and *B*2 100 units at a unit cost of 6. There are two entrepreneurs, *E*1 and *E*2. *E*1 procures only from *B*1 and *E*2 from *B*2. They incur no access costs. The dashed vertical line indicates that *A*1, *B*1 and *E*1 are co-located, as are *A*2, *B*2 and *E*2.

The only connections are those established by entrepreneurs *E*1 and *E*2. *E*1 can serve only *A*1 but *E*2 can serve both *A*1 and *A*2. For simplicity it is assumed that both entrepreneurs have integrated backwards and own their own production facility. For simplicity all access costs are set to zero.

Solution of the model

Because there are two entrepreneurs who are rivals in serving market *A*1 the retail price in *A*1 will be bid down to *B*2's unit cost of 6. However, *E*2 has a monopoly of *A*2 and so the retail price in *A*2 will be fixed at the customers' reservation price of 8.

*E*1 monopsonises *B*1 and *E*2 monopsonises *B*2; hence the wholesale price of *B*1 is equal to the unit cost of 4 and the wholesale price of *B*2 is equal to the unit cost of 6.

*E*1 therefore supplies 40 units from *B*1 to *A*1, and *E*2 supplies 20 units from *B*2 to *A*1. *E*2 also supplies 60 units to *A*2.

The distribution of income

*E*1 earns a margin 6 – 4 = 2 on 40 units, giving a profit of 2 x 40 = 80. This reflects their exclusive access to the cheapest source of supply *B*1. *E*2 only breaks even on exports to *A*1 but earns a margin of 8 – 6 = 2 from its domestic market *A*2 because of its monopoly. Thus *E*2 makes a profit 60 x 2 = 120. Consumers in *A*1 earn a unit surplus of 10 – 6 = 4 on 40 units purchased, giving a total surplus 4 x 60 = 240. Total value created is (10 x 60) + (8 x 60) = 1080. Total cost is (4 x 40) + (6 x 80) = 640, and total surplus is 1080 – 640 = 440, which is divided up as above.

Linear programming solution

Because of the simple nature of the example above it was possible to derive the solution of the model by plausible reasoning. This becomes very difficult as the numbers of customer segments, producers and entrepreneurs increase, competition intensifies and linkages proliferate. It is sometimes possible to 'guess' solutions by comparing the initial profitability of different linkages, as suggested in Casson (2024), but this is not a reliable method. Linear programming is therefore the appropriate method of solution, as it will solve any model of the general type considered here. The number of dimensions to the analysis is equal to the number of linkages available.

When expressed as a linear programming (LP) problem the maximisation of the gains from trade subject to demand and resource constraints appears as follows:

Maximise $v = 10(x_1 + x_2) + 8x_3 - 4x_1 - 6(x_2 + x_3) = 6x_1 + 4x_2 + 2x_3$

subject to:

$$x_1 + x_2 \leq 60 \quad [A1\text{'s satiation demand}]$$

$$x_1 \leq 40 \quad [B1\text{'s production capacity}]$$

$$x_2 + x_3 \leq 100 \quad [B2\text{'s production capacity}]$$

$$x_3 \leq 60 \quad [A2\text{'s satiation demand}]$$

giving the same solution as presented above, namely

$$x_1 = 40; \quad x_2 = 20; \quad x_3 = 60.$$

Diagrammatic analysis

Where there are just two or three potential linkages it is possible to illustrate the simplex method geometrically. The two-dimension case can be illustrated using a simple diagram with horizontal and vertical axes, while the three-dimensional case can be solved using a two-dimensional projection of three-dimensions shapes.

The solution of the problem above is illustrated in Figure 3. The diagram solves for quantities but not for prices. Each axis represents the volume of trade along one of the linkages. The linkage from $B1$ to $A1$ through $E1$ is represented by the horizontal axis x_1, the linkage from $B2$ to $A2$ through $E2$ by the axis horizontal x_2, which lies at right-angle to the axis x_1 along the 'floor' of the diagram, while the linkage from $B2$ to $A1$ through $E2$ is represented by the vertical axis x_3.

FIGURE 3 HERE

The multi-faceted box shown in the figure is a 'convex hull'. It represents the set of feasible outcomes. The most productive outcomes lie on the surface of the hull. The hull is said to be convex because any straight line drawn between any two points inside or on the surface of the hull will also lie entirely inside or on the surface of the hull. The number of available linkages determines the dimension of the hull. In the present case there are three available linages and so the hull is three-dimensional. The figure portrays a two-dimensional projection of this three-dimensional body.

The maximand is the gains from trade, which is the linear function v, as presented above. This linear function defines a family of planes which move away from the origin (where the value of output is zero because nothing is produced) and towards the reader as the value of v is increased. The planes are not shown because they would unduly complicate the figure. As the plane moves outwards and upwards from the origin of the figure it increases in value. The highest attainable value lies at the point where the ascending plane is tangent to (i.e. just touching) the surface of the hull. This point is identified by a black dot in the figure. The coordinates of this point, when referred to the three axes, determine the equilibrium quantities traded along each of the linkages. These quantities in turn determine the amounts produced by each producer and the amounts consumed by each customer segment.

The attraction of the LP approach is that it can be extended to cope with many customer segments, many producers and many possible linkages between them, intermediated by multiple entrepreneurs. The simplex method, as its name suggests, provides a simple and efficient algorithm for solving LP problems, and it is readily available free of charge as part of a laptop utility.

PART THREE: *CASE STUDIES*

Three special cases are now presented, each illustrating a different application of the analysis. Each case examines the dynamics of prices, quantities, profits and market shares over time. Readers who have studied Part Two will recognise these case studies as applications of the model of segmented markets presented above. Readers who have not studied Part Two will learn much about the model from these applications.

The examples below are designed to be as simple as possible, but because of the complexity of the issues with which they engage, they require very careful reading. Readers interested only in general principles may therefore prefer to proceed directly to Part Four.

Case 1: Evolution of a market through sequential entry

In this example three entrepreneurs evaluate entry in sequence, beginning with $E1$ and ending with $E3$. After the final stage, the product becomes obsolete.

Stage 1. Entrepreneur $E1$ is the first to recognise a potential opportunity and, after appraisal, decides to enter. They believe, correctly, that customer group $A1$ has a demand for 100 units of a novel product at a reservation price of 12, but they can only identify one producer, $B1$. They believe, correctly, that $B1$ has a unit cost of 5, and they offer them a take-it-or-leave-it price of 5 which $B1$ accepts. However, $B1$ has a capacity of only 60 units, while $A1$ demands 100 (although only $E1$ knows this). Thus $E1$ supplies 60 units to $A1$ at a price of 12, making a profit margin $12 - 5 = 7$ and a total trading profit of $7 \times 60 = 420$.

$E1$ is aware, however, that another entrepreneur might enter and bid against them for $B1$'s output. They therefore acquire $B1$ (e.g. through take-over) in order to pre-empt any potential entrant, as indicated by the dashed box on the left-hand side in Figure 4. This means that potential supply from $B1$ to a later entrant is blocked.

FIGURE 4 HERE

Stage 2. Sure enough, entrepreneur $E2$ becomes aware that there is unsatisfied demand for the product and decides to enter the market, knowing the situation at the end of stage 1. They cannot gain access to $B1$, but they can gain access to $B2$, of which $E1$ is unaware. $B2$ has unit costs of 6 and a capacity of 60. $E1$ and $E2$ now have a combined capacity of 120, which means that there is excess supply in $A1$. Competition between $E1$ and $E2$ now bids down the price to 6. $E1$ prices at just below 6 and sells 60 units (as before) while $E2$ prices at 6 and sells the residual 40 units. $E1$ earns a profit margin of $6 - 5 = 1$ and makes a profit $1 \times 60 = 60$, whilst $E2$ just breaks even. Customers in $A1$ receive a total surplus of $(12 - 6) \times 100 = 600$.

The best strategy for $E2$ would be to collude with $E1$. For example, by forming a cartel or by merging they could increase the price from 6 to 12 and increase their combined profit by $6 \times 100 = 600$. $E2$ could in principle bargain for the whole of this increase in profit, but in practice they may have to settle for less, e.g. some agreed proportion of the total profit, which is equal to $600 + 60 = 660$. They could agree to split the difference, although $E1$ might demand 60 per cent on the basis that that was the share of product that they supplied to the market.

Stage 3. Potential entrant $E3$ studies the situation at the end of stage 2 and resolves to build a new plant on a larger scale that will have lower unit costs. The new plant has a capacity of 120 and unit costs of 3. $E3$ realises that $E1$ and $E2$ must be barred from accessing these supplies and so they ensure that $B3$ is a wholly-owned subsidiary (illustrated by the dashed box). $E3$ now has a choice. They can price at just below 6 and share the market with $E1$ or they can price at just below 5 and control the entire market themselves. If they share the market equally (through collusion) then they earn a profit of $(6 - 3) \times 50 = 150$ and $E1$ makes a profit $(6 - 5) = 50$. If they price at just below 5 (i.e. competitively) then they make a profit $120 \times (5 - 3) = 240$, which is larger. Since this is larger, they price at 5 and take the whole market. Note that $E3$ becomes a monopolist because their price is low, which is contrary to the predictions of textbook monopoly theory, which suggests that monopoly will cause the price to be high. The explanation is that it is the threat of competition rather than actual competition that keeps it low. This threat, in turn, arises from the exploitation of economies of scale.

Further refinements could be added to the model. For example, if the novelty of the product wore off and demand was halved then, if $E3$ had large recurrent overhead costs (exceeding 60 x 2 = 120) that could be recovered, they would choose to leave the market. The market would return to $E1$ and $E2$ and the price would go back up to 6.

The impact of these changes on the entrepreneurs depends on how long each of these stages lasts. If entry is rapid then the total income generated by $E1$ and $E2$ may be quite small. If each stage lasted one year then $E1$ would earn an annual profit of 420 in the first year, 60 in the second year and nothing in the third year. Whatever the lengths of the periods, $E2$ earns no profit at all, because they rely on the highest-cost producer in a market with excess capacity. However, they have a significant role in determining what profit the others make. They are the 'customer's friend', because they help to keep prices low by giving customers a choice that they otherwise would not have. They also face a 'moral dilemma', because they could bargain for a substantial sum if they agreed to collude or sell out.

The time scale involved may be very different from that assumed above, however. Nothing will alter the fact that $E2$ can never make a profit except by doing a deal, but the time scale has a big impact on whether the first entrant or the last entrant makes the most profit. The point is obvious, but is sometimes overlooked, because both popular sentiment and academic research both seem to prioritise the innovator, or first-mover, over the later entrants, even though the later entrants typically dominate larger markets, control more complex organizations, and, because of their size, make larger profits.

Innovation of a technology

In this example two entrepreneurs, $E1$ and $E2$, compete to exploit a new proprietary technology owned by entrepreneur $E0$. $E0$ has employed researchers $R0$ to generate the technology, which they have patented (see Figure 5). There is no limit to the quantity of product that the new technology can be used to generate, other than the availability of suitably qualified labour and the finite demand for the product they produce.

FIGURE 5 HERE

Stage 1. Initially the product is procured without access to the new technology. The technology is then developed at a cumulative cost of 800, which must be recovered during the single long period over which it is exploited. At the end of this period the technology becomes obsolete.

Customer valuation of the product remains unchanged after the new technology is introduced, but demand increases. There are two segments of demand ; the reservation price in segment $A1$ is 15, and in $A2$ it is 12. Initially 100 units are demanded in each segment. Both entrepreneurs $E1$, $E2$ have access to both segments. $E1$ relies exclusively on producer $B1$ who, in the absence of the new technology, has unit costs of 9, while $E2$ relies on $B2$, who has the same capacity, namely 120, but a marginally higher unit cost of 10. Competition between $E1$ and $E2$ bids down the price in both markets to 10. $E1$, having the lower cost, prices just below 10 and sells all their potential output, 120 units, leaving, $E2$, pricing at 10, to supply the residual 200 - 120 = 80 units. $E1$ makes a profit $(10 – 8)$ x 120 = 240 and $E2$ merely breaks even. Customers in segment 1 receive a surplus $(15 – 10)$ x 100 = 500, and customers in segment 2 a surplus $(12 – 10)$ x 100 = 200.

Stage 2. *E*0 now enters, having developed a new technology at a cost of 800. The new technology improves the productivity of the workforce, reducing unit cost in *B*1 from 9 to 5, and reducing unit cost in *B*2 from 10 to 6. Higher productivity per worker also increases the potential output of both producers from 120 to 200. The new technology also makes the product more versatile, thereby increasing demand from 100 to 150 in both markets.

*E*0 is aware of the pre-innovation scenario and recognises that competition between *E*1 and *E*2 will potentially bid down customer prices. They therefore exploit their monopoly of the new technology to sell *E*1 sole rights to supply segment *A*1 and *E*2 sole rights to supply segment *A*2. If they know the prices at which *E*1 and *E*2 can procure then they can set the cost of the licences they sell to *E*1 and *E*2 in order to maximise the profit they can appropriate for themselves. *E*1 and *E*2 do not have to adopt the technology, however, and so *E*3 will allow them to retain all the profit they received before. In this case *E*1 will sell 150 units to *A*1 at a price of 15, generating a revenue 2250. They will incur production costs of 5 x 200 = 1000 and will retain their previous profit of 240. Since both *E*1 and *E*2 know that *E*0 can, if necessary 'play them off against each other', they do not bargain for more than their previous profit. *E*0 therefore makes a profit of 2250 – 1000 – 240 = 1010 from their licensing to *E*1. Likewise they can license to *E*2 the right to supply 150 units to *A*2. In the absence of competition from *E*1 this will generate a revenue 12 x 150 = 1800. Production costs will be 6 x 150 = 900. Because of competition under the previous regime, *E*2 makes no profit. *E*0 can therefore demand a licence fee of 1800 – 900 = 900. *E*0 therefore generates a total revenue of 1010 + 900 = 1910 and a net profit of 1910 – 800 = 1110.

Note that *E*0 could, in principle, earn a modest additional profit of 50 if they licensed *E*1 to sell 200 units and *E*2 only 100 units, because *E*1 could then supply 50 units to *A*2, produced by *B*1 at a unit cost of 5 rather than by *B*2 at a unit cost of 6. This would mean, however, that they would have to police rivalry between *E*1 and *E*2 in *A*2, which could potentially cost more than the savings achieved.

Choice of production technique

In this example (see Figure 6) the entrepreneur *E*0 has obtained a monopoly of a mineral with a maximum extraction rate of 4000 units per period. The unit cost of extraction is 2. The mineral needs to be refined and processed before the finished product can be sold to the customer group *A*1. This group demands 1000 units of final product per period at a reservation price of 20. There are two production plants that can process the mineral but they use different technologies. Producer *B*1 combines 2 units of labour supplied by labour force *L*1 with four units of the mineral to produce one unit of finished product, while *B*2 is more labour-intensive and less mineral-intensive, and combines 4 units of labour from labour force *L*2 with 3 units of mineral to produce one unit of finished product. The price of labour to *B*1 is 5 per unit and the price to *B*2 is 3.5. *B*1 can employ up to 1200 workers and *B*2 up to 1500 workers.

FIGURE 6 HERE

The key decisions are taken by the entrepreneur *E*0, who has all the information they require to maximise their profit. The mine is operated by producer *M*0, which is a subsidiary of *E*0. *E*0 uses the downstream retail entrepreneurs *E*1, *E*2 to distribute their product. These entrepreneurs have integrated backwards into production; *E*1 owns and controls *B*1 and *E*2

owns and controls $B2$. $E0$ knows the input-output coefficients for each production process. Given these technologies, $E0$ has access to sufficient mineral output to satisfy the global demand for 1000 units of processed mineral per period.

The maximum retail price that $E1$ and $E2$ can charge is set by the reservation price in the global market, which is 20. For each unit supplied by $B1$ to $E1$ the unit labour cost is 2 x 5 = 10, and the amount of raw mineral required is 4. $B1$ is locked into supplying $E1$, and so the profit margin per unit of raw mineral appropriated by $E1$ is (20 – 10)/4 = 2.5. For each unit supplied by $B2$ to $E2$ the corresponding profit margin per unit is (20 – 14)/3 = 2. To maximise profit $E0$ will sell as much output as possible to the entrepreneur who can afford to pay the higher price. This is $E1$, whose profit margin is 2.5. By working at full capacity $B1$ can produce 1200/2 = 600 units, leaving 400 units to be supplied by $B2$. $B1$ therefore uses all the 1200 units of labour available to them, while $B2$ therefore only 3 x 400 = 1200 units of labour, which is less than the 1500 available. $E0$ supplies 600 x 4 = 2400 units of mineral to $E1$ for use by $B1$, and 400 x 3 = 1200 units to $E2$ for use by $B2$, leaving spare capacity in mineral output of 4000 – 2400 – 1200 = 400. Neither $E1$ nor $E2$ appropriates any profit because of the monopsony power of $E0$. $E0$ makes a profit of (2.5 x 600) + (2 x 400) = 2300.

Sequential games: anticipating rivals' reactions

It has so far been assumed that each decision taken by an entrepreneur is essentially myopic. Entrepreneurs achieve flexibility because they have subsequent opportunities to modify their earlier myopic decisions. It is possible, however, to analyse the competitive process in a different way, in which entrepreneurs are sufficiently far-sighted to correctly predict the entry of subsequent entrepreneurs and the reactions of incumbent entrepreneurs to such events. In this case the behaviour of the entrepreneurs can be analysed as the outcome of a sequential non-cooperative game [Bloch and Ferrer, 2001; Faias, Herves-Beloso and Moreno-Garcia, 2010; Shapley and Shubik, 1969; Toraubully, 2018].

Given the judgemental nature of entrepreneurship, and the lack of perfect foresight, it is absurd to suppose that in practice this approach could predict the actual behaviour of entrepreneurs. The exercise is, however, relevant to the analysis of entrepreneurship because it provides a standard of decision-making to which an ambitious entrepreneur could aspire. Furthermore, in an era of computers and artificial intelligence, it represents the kind of 'decision support' that could be offered commercially to an ambitious entrepreneur. Using the software, an entrepreneur could generate a range of possible scenarios and then use a market model to optimise their response to each scenario. The strategy that produced the best overall results in the simulations would then be selected.

The structure of a sequential game extends the analysis of the scenarios discussed above to the case of any market in which entrepreneurs make sequential moves. For the purposes of the model the sequence of moves is fixed from the outset. Each entrepreneur knows not only the state of the market at the time they enter but the potential for further entry later. This implies that the first-mover must be very far-sighted as they (alone) know the entire sequence of moves that will take place. For other aspects of competitive strategy see Clark, 1940, 1961 and Porter, 1980.

At their time of entry each entrepreneur correctly predicts the moves that other players will make, including the moves of players who have not yet entered the market. Each entrepreneur

attempts to maximize their profit in full possession of all the information they require about the entry process. This information includes the profits that will accrue to the other entrepreneurs under each possible set of outcomes. This information allows each entrepreneur to predict the decisions of all the other entrepreneurs, conditional on their own choice of strategy.

Sequential games are solved recursively, with the last mover deciding what they will do conditional on each possible state that is a legacy of earlier entry decisions. The second-last mover then decides what they will do, conditional on previous decisions and their correct prediction of the final mover's response to the situation they face. This process is repeated until the first-mover has made their decision, at which point the entire sequence of entry is determined.

Once the game has been solved it is possible to calculate the discounted value of the rewards accruing to each of the entrepreneurs who enters the market, valued at the time they make their initial entry. However, each entrepreneur may have a different rate of discount; this may reflect the impatience of the entrepreneur for early profits, and possibly their perception of risk. In the context of the model there is no risk, since each entrepreneur correctly predicts the behaviour of the others, but the entrepreneurs do not necessarily know this, and so some may perceive the risk to be high.

The players of a sequential game are far-sighted, as explained above. For example, a first-mover's strategy may, from the outset, be dominated by concerns regarding long-term market share. This suggests, for example, that a first-mover may sacrifice substantial short-term profits to discourage entry, e.g. by selling at a low price from the outset, or by reinvesting short-term profits in long-term technological improvements or in strengthening brand identity.

Summary of the case studies

In this analysis of market evolution the emphasis has been on the discovery and diffusion of information, and the way that entrepreneurs react to each other when they compete. Market evolution can be a lengthy process, and many exogenous changes can occur as a market matures. These case studies have shown that the entry process is extremely complicated, so that, even in the absence of exogenous changes, the evolution of any given market will be hard for anyone, including those involved, to predict.

The cases examined above are of considerable practical significance. While the models presented above cannot predict the future they may be useful in retrodiction. They can be used in historical case studies to interpret the thinking that must have underpinned the behaviour of entrepreneurs during the evolution of markets in the past. This would allow key historical steps in market evolution to be understood in terms of the entrepreneurial thinking that may have been used at the time that key decisions were made.

These examples of market evolution illustrate very clearly how the income accruing to any particular group – whether consumers, producers or entrepreneurs – is subject to almost continual disturbance as circumstances change. There are many factors that can change and many transmission mechanisms that can propagate a given change from one segment of a market to another.

Diversity of applications

The examples have highlighted the very considerable degree of complexity which the theory can engage.

It is possible to vary the number of potential customer segments, the number of different potential producers and the number of different routes by which customers may be linked to producers. Entrepreneurs can select different methods of linkage, e.g. different modes of transport, different inventory management methods, and so on. There may also be different technologies available for production. Some technologies may be known to some producers but not to others.

For any given technology production may involve combining different types of skilled labour in a distinctive way, while locations may differ in the types of skill with which they are endowed. For simplicity, however, it must be assumed that production capacity is determined by the managerial capacity of the producer and not by a local shortage of any specific skill.

Each of these changes complicates the analysis but all of these complications can be readily handled using the methods described above. All that happens is that each linkage specifies not only the customer segment, the entrepreneur and the producer, but also the transport route employed, and technology used. In addition, the choice of producer identifies the local labour market from which they recruit. Each producer has, by assumption, monopsony power in their labour market, so that they need to pay each worker no more than their opportunity earnings, as determined by their productivity in producing the basic product for themselves.

Remarks on the case studies

In the examples above the changes were driven mainly by the acquisition of information and the consequent reactions of entrepreneurs - in particular, their discovery of new sources of demand and new locations of production, and the potential that they offered. The drivers were not exogenous shocks in the sense of changes in the physical environment or changes in technological know-how. Nothing actually happened to instigate the process apart from a changes of perception.

As each entrepreneur made some discovery they introduced a competitive element which obliged other entrepreneurs to react. In order to mitigate a threat created by another entrepreneur, each entrepreneur had to react, and that reaction itself created threats for other entrepreneurs. The drivers of change were therefore not only the acquisition of information by an entrepreneur, but the reaction of other entrepreneurs to their response. All this happened without any change in the original circumstances. The key activities of the entrepreneurs were the discovery of new locations, gaining access to those locations, and negotiating with the people there.

In the absence of entrepreneurs, therefore, an economy may be quite rigid. It may be self-sufficient, and therefore insulated from changes elsewhere in the global economy. Intermittent changes may, however, have serious effects on specific parts of the economy, because there is limited scope for internal adjustment. An entrepreneurial economy, on the other hand, may experience more changes, because it is better integrated into the global economy. Its response to these changes will be more flexible, and so their impact will be diffused throughout the economy.

PART FOUR: REFLECTIONS ON THE MODEL

Specialisation and the division of labour

A simple way to understand the emergence of the different roles identified above – customer, producer, retailer, wholesaler, entrepreneur – is to recall Adam Smith's theory of specialisation by the division of labour [Smith, 1776]. Smith argued that a 'jack of all trades' will have low productivity, but a specialist will have much higher productivity because they can 'learn by doing'. The more specialised the task, the quicker it is to perform, and the more frequently it is repeated, the higher the productivity that can be achieved. Everyone would therefore benefit from specialisation. When everyone was specialised they would each produce one thing but want many things, and therefore they would need to acquire the other things that were missing. The easiest way for them to do this would be to attend a market where they sold their own product and then used their income to buy the other products that they needed from other traders at the market.

Smith did not take the next step, however, which is to argue that market-making is itself a specialised activity. Although he referred to the 'butcher' and the 'baker', he did not explicitly identify these as specialised activities. The specialist producer will sell to the specialist merchant who will attend the market to sell the goods. In this context the merchant is an entrepreneur who buys in bulk, as a wholesaler, at the farm gate or the factory door, and sells in smaller units in an urban market place. The market can therefore be open whilst the producer is still at home, on the farm or in the factory, producing the good in which they specialise. The importance of merchants throughout recorded history (they are frequently mentioned in the Old Testament, for example) testifies to the historical significance of the entrepreneur as retailer.

Note that while retail markets are spatially concentrated, as merchants agglomerate in the market place, early wholesale markets were dispersed, as itinerant merchants visited producers to collect supplies. As the costs of transport and communication diminished over time, however, wholesale markets became concentrated in financial centres, where claims on goods were traded, and the goods themselves were then shipped direct to buyers by road or rail. The arrival of the internet and the emergence of 'platform firms', has further stream-lined marketing and distribution systems [Cusumano, Gower and Yoffie, 2019].

Comparative advantage, gains from trade and Pareto efficiency

To appreciate the significance of the economic principles embedded in the models presented above it is essential to understand how these principles evolved. The key principles are comparative advantage, the gains from trade and Pareto efficiency.

Although Adam Smith believed that specialisation increased productivity, he also believed that anyone could specialise equally well in almost anything. However his follower, David Ricardo, believed that, on the contrary, people and places possessed natural aptitudes, or natural endowments, which gave them 'comparative advantage' in specific industries and occupations. People would specialise in the activities in which they had comparative advantage and use their earnings to purchase goods from other people whose comparative advantage lay in producing other products that they wanted but could not readily produce themselves. Comparative advantage was different from absolute advantage, where a person or place might be better than another at everything. Comparative advantage determined the

pattern of specialisation, while absolute advantage determined differences in the level of income earned in a given occupation [Ricardo, 1817].

John Stuart Mill developed Ricardo's approach and attempted to determine the terms of trade between groups of people who chose to specialise in producing different goods, both of which everyone liked to consume. These terms of trade were expressed in the relative price of the two goods. The price was measured in units of one of the goods, and so varied depending on which good was used for this purpose [Mill, 1844].

The problem of price determination turned out to be more difficult than Mill had imagined. Mill devised 'offer curves', one for the buyers and one for the sellers, and argued that price and quantity were set where these curves intersected. Francis Ysidro Edgeworth, however, noted that buyers could make themselves better off by forcing the seller onto a different part of their offer curve, and that the converse also applied to sellers. Edgeworth devised a 'contract curve' along which each party optimised their response to the other, but he could not determine where on the contract curve an agreement would be reached. To simplify his exposition Edgeworth assumed that people exchanged endowments rather than products, and traded as individuals rather than groups. This abstraction reduced the practical relevance of his theory and inhibited the diffusion of his ideas [Creedy, 1986; Edgeworth, 1881, 1925; Walker, 1973].

Marshall, seeking to reach a popular audience, decided to focus on a market for a single product and to collapse all other markets into some other market which handled everything else [Marshall, 1890]. This became known as 'partial equilibrium analysis'. In this market the product was exchanged for money, and the money was then spent in the other catch-all market. Marshall then measured the gains from trade in terms of money. Money was the proxy for all the other goods that people could produce and consume, i.e. it was the 'basic product' in the model presented above.

Understanding the limitations of Marshall's analysis is the key to developing the novel theory of the entrepreneurship set out above. Under the influence of Mill and Edgeworth, Marshall emphasised the adversarial aspect of trade, namely the bilateral negotiation of price. Each buyer wanted to buy cheap and each seller wanted to sell dear. Each buyer was looking for the cheapest seller, while each seller was looking for the buyer willing to pay them most. Marshall realised that it was a complicated problem, of a type that Edgeworth previously had failed to solve [Casson, 2003; Sutton, 2002; Creedy, 1992, 1998].

Marshall was aware that the French economist Leon Walras had already proposed a fictional auctioneer, who would take bids simultaneously for each unit of the product that could be produced or consumed [Walras, 1954]. Marshall wanted his account to be realistic, however and so he simply fudged the issue of price determination by referring to 'market tendencies'. He argued that the force of demand, driven by the self-interest of buyers (wanting to satisfy their wants at the lowest possible price), would neutralise the force of supply, driven by the self-interest of sellers (wanting compensation for the increasing marginal cost of working close to capacity). Balance would be achieved at a point of equilibrium where the supply and demand schedules intersected. The height of each schedule would reflect the slope of the relevant offer curve in Mill's analysis. Uniqueness and stability were guaranteed so long as the demand schedule cut the supply schedule only once, and, as quantity increased, cut it from above.

This populist appeal to forces was based on mechanical and electrical analogies that were fashionable at the time. Mechanical forces derived from gravitation and inertia, while electrical forces depended on attraction and repulsion, neither of which has any direct analogue in economics. Marshall's forces may, however, be interpreted as representing self-interest. The self-interest of the buyers is pitched against the self-interest of the sellers. The two forces battle with each other until a stalemate is reached, which represents the equilibrium. Thus price is the battlefield, negotiation is the weapon and equilibrium is the outcome.

Implications for entrepreneurship

Despite appearances to the contrary, there are two key points in the preceding analysis that point directly to a theory of the entrepreneur.

Firstly, Marshall's equilibrium not only equates the 'forces' of demand and supply, but also maximises the gains from trade. In a supply and demand diagram the gains from trade are measured by the area underneath the demand schedule but above the supply schedule, identified by Marshall as the sum of consumer surplus and producer surplus. The gain accruing to consumers from any given unit of production is measured by the difference between the equilibrium price and the height of the demand schedule, and the gain accruing to producers is equal to the difference between the equilibrium price and the height of the supply schedule. Thus the overall gain from trade in a given unit is equal to the difference between the height of the demand schedule and the height of the supply schedule at the appropriate point on the horizontal axis. The total gain from trade is therefore the sum of these across all the units traded, i.e. the area between the schedules to the left of the equilibrium.

Secondly, if an entrepreneur had superior information, so that only they knew what the consumers demanded and what the producers were able to supply, then they could, in principle, appropriate all the gains from trade for themselves. They would discriminate in price, charging each customer for each unit the maximum that the customer was willing to pay, and paying each seller for each unit the minimum that that seller was willing to accept. The profit of the entrepreneur would equal gains from trade, and *vice versa*. This prospective profit would be a powerful motivation for some individual to specialise in searching out this opportunity by committing to becoming an entrepreneur. From this perspective, the entrepreneur is like an arbitrator, or peacemaker, on the battlefield described above. They settle disputes by appropriating all the gains for themselves.

There remained two outstanding issues, however. The first is that to maximise the gains from trade when demand decreases as price increases the entrepreneur must discriminate between individual customers if they are to appropriate all the gains from trade; likewise the entrepreneur must discriminate between producers if they are reliant on multiple producers to satisfy demand. These requirements can be eliminated, however, if all the consumers demand a specific amount of product at the same reservation price and all producers have the same unit costs. These conditions generate, in total, a rectangular demand schedule and rectangular supply schedule, and so discrimination between customers, or between suppliers, is unnecessary.

The interaction between a rectangular demand schedule and a rectangular supply schedule generates one of two scenarios; in the first there is excess supply of the product, and the demand constraint is binding on the entrepreneur, and in the second there is excess demand and the supply constraint is binding instead. As a result, output is equal to the minimum of demand and supply and the profit on each unit supplied is equal to the excess of the reservation price over the unit cost of production. This is the scenario used in the exposition of the theory presented in Part Two.

This analysis implies that if demand and supply schedules are rectangular then there is no loss of the gains from trade from the variation of price within the limits set by the customers' reservation price and the producers' unit cost. This in turn indicates that if an entrepreneur charges a uniform monopoly price to a group of customers then there is no reduction in the gains from trade; there is merely a redistribution of these gains from the customers to the entrepreneur. Likewise, if an entrepreneur purchases at a uniform monopsony price from a producer there is merely a redistribution of gains from the producers to the entrepreneur. Thus the exercise of monopoly or monopsony power by an entrepreneur has no impact on the gains from trade. The market process maximises the gains from trade, not only under competition, but also when there is a monopolistic or monopsonistic entrepreneur. This result provides a bridge from Marshall's theory of the market to the theory of entrepreneurship that was presented in Part Two above.

The second outstanding issue is that there is only one entrepreneur in this account of the market. If there were two or more entrepreneurs then they would bid up the supply price and bid down the demand price until all profit was eliminated. This predicts correctly that the prospect of competition will discourage entry. If entry occurs, and there is excess demand, then the wholesale price will be bid up to the customer reservation price and all the profit will accrue to the producers, while if there is excess supply then the retail price will be bid down to the unit cost of production and all the gains will accrue to the customers. This principle was incorporated in pricing rules that were applied to the model presented in Part Two.

This analysis also shows that the introduction of the entrepreneur eliminates the 'indeterminacy' of bilateral monopoly, which was a prominent feature of post-Marshallian market theory [Hicks, 1932]. Under the assumed conditions a monopolistic entrepreneur will simply use their information advantage to buy at the producer's unit cost and sell at the consumers' reservation price, and thereby appropriate all the profit for themselves.

Comparison with modern neoclasssical theory

In view of the radical nature of the theory proposed above, it is useful consider whether the concepts introduced in this new approach to entrepreneurship theory contains any analogues of concepts already familiar from conventional theory. The answer is that they do. The reader may have noted that there has been no explicit reference to indifference curves, production possibility curves, and other concepts from neoclassical economics. However, these concepts have analogues within the model set out in this paper. The answer, in brief, is that the missing 'curves' are replaced by kinked straight lines, but these kinked straight lines are redundant for the exposition of the theory. The theory is at once much simpler and more sophisticated than neoclassical theory. The proliferation of redundant concepts in neoclassical theory simply makes that theory appear to be more sophisticated than it really is.

The *indifference curve*. The analogue of a consumer indifference curve is an iso-value line. When consumption of the novel product is measured along the horizontal axis and consumption of the basic product along the vertical axis then the iso-value line is a kinked straight line. To the left of the kink it is a backward-sloping straight line with a slope equal to the reservation price of the novel product, and to the right of the kink it is horizontal. The steepness of the slope indicates the amount of the basic product that must be sacrificed in order to increase consumption of the novel product by one unit. The slope becomes zero once satiation consumption of the novel product has been reached.

The *production possibility curve*. This curve is widely used in international trade theory to analyse specialization between two countries in the production of two products. In the present context the two products are the standard product and the novel product. The potential output of each product is plotted along an axis. The two counties specialise according to comparative advantage. When worker productivities differ between the two countries one country specialises in production of the basic product and the other in production of the novel product. Eventually a point is reached where all the workers in one country are fully specialised in the production of one of the products, so that some of this product has to be produced in the country without the comparative advantage, thereby increasing its unit cost. This point of transition generates a kink at the meeting point of two straight lines with differing slopes. Just as with the indifference curve, the production possibility curve is transformed into a kinked straight line.

The solution of the entrepreneurship model does not require the construction of either of these lines, since the solution algorithms discussed are perfectly adequate for the purpose. Where appropriate, however, the solution can be portrayed using the kinked straight lines that approximate to these curves. Just as in mainstream theory, the solution lies at a point of tangency between these two lines. Unlike mainstream theory, however, the point of tangency is at a kink in one of the lines, which means that the slope of the tangent is not unique. The difference in slope between the steepest tangent and the flattest tangent measures the mark-up on product that can be appropriated by an entrepreneur that exercises both monopoly and monopsony power.

Similar remarks can be made about the 'isoquant' and the 'production function'. An isoquant shows the trade-off between alternative combinations of complementary factors of production that produce the same amount of a given product. In the context of the models presented above, an isoquant is most conveniently derived by postulating two or more alternative techniques of production, which combine two or more types of productive resource in different fixed proportions under constant returns to scale. A simple example would involve two labour intensive-production techniques that combined two given types of skilled labour in different proportions. In these circumstances the isoquant is simply a kinked straight line, convex to the origin, with the kinks positioned at the points of intersection of the line with rays, emanating from the origin, whose slopes are equal to the proportions in which each technology combines the two respective types of labour. A production function can be readily derived from a family of isoquants, corresponding to different amounts of output, by simply postulating constant returns to scale with respect to the amount of labour used.

Price-cost mark-up. In mainstream neoclassical theory the price-cost mark-up is normally discussed only in the context of monopoly power. Empirical studies consistently show that

these mark-ups vary in size between firms in the same market, and are often large. This is exactly what is expected in a segmented market. It is a problem for mainstream theory, however, because mainstream theory uses only a measure of average market power. Mainstream theorists typically suggest that the more firms there are, the more competitive the market will be, and so the smaller the mark-up will be. Though plausible, the reasoning is loose and conjectural by the normal standards of mainstream theory. The theory of market segmentation, by contrast, explains why the opposite may occur. This is because a market with many isolated segments may have many local monopolies, and hence many firms each with a high mark-up; by contrast, a market with only a few segments, each served by the same few entrepreneurs, may be much more competitive and therefore have lower mark-ups.

Comparison with Austrian theory

Some readers who are strongly committed to Austrian theory may feel that this paper simply offers an extension of, rather than a radical alternative to, Austrian theory. They may argue that everything described above is ultimately reducible to arbitrage and that this paper has simply complicated the analysis rather than clarified it. This would be a mistake, however.

The present theory is not primarily concerned with pure arbitrage, where the product has already been produced and is then re-sold. In the models described above the entrepreneur does not identify someone who has already acquired a product to which they attach a modest value, and pair them up with someone else who values the product more. Instead they identify a customer segment where there is an unsatisfied potential demand for a product for which there is as yet no adequate supply. Having identified a customer segment, they then identify a producer who has the potential to supply the product in question at a price less than the price that they believe potential customers are willing to pay. They then establish a linkage between the source of demand and the source of supply over which they have exclusive control. They buy from the producer and re-sell to the customer, recovering their costs out of the gross profit margin between buying and selling price.

If there is no other linkage of the same kind then the entrepreneur will acquire monopoly power. To protect this monopoly they will discourage rival producers from entering the customer segment, and discourage their customers from seeking out rival producers.

When markets are completely segmented, information about the novel product will not diffuse through normal social channels, so that other people who could in principle benefit from it will not get know about it too. People in other segments may either have different requirements, or may not be in touch with anyone in the segment that is already being served.

For a product that has been successfully innovated in a single segment to become more widely used, therefore, potential customers in other segments must be identified either by the innovator (the 'first-mover'), or by some other entrepreneur (a 'second-mover'). The second-mover may be either an imitator or an originator, i.e. they may be alerted by the success of the first-mover or they may discover the opportunity independently for themselves.

If the first-mover completely fails to recognise the potential for a wider market then they will not seek to enter the second market once it has been discovered, and so the second mover will acquire a monopoly. This monopoly may be transitory however, as other entrepreneurs may observe their success and attempt to enter too.

To summarise, while segmentation may be said to provide opportunities for arbitrage, it is not the kind of arbitrage commonly described in connection with commodity markets or financial markets, and not the kind that is emphasised in popular expositions of Austrian theory. In these cases market institutions already exist and the entrepreneur simply buys and resells in either the same market or in some other pre-existing market. In the case of the entrepreneurs described above, however, there may be no pre-existing market. They create the market in which they trade by establishing contact with people who have not yet indicated a willingness to trade.

It also needs to be recognised that while arbitrage can reduce price dispersion, it cannot normally eliminate it altogether, because the arbitrageur incurs costs which drive a wedge between their buying price and selling price. Barriers to arbitrage include the opportunity costs of monitoring prices, the transactions costs incurred in negotiating and enforcing contracts, and, where speculation is involved, the costs of storage. The theory of entrepreneurship presented in this paper therefore offers a more nuanced account of arbitrage and speculation than that provided by conventional Austrian theory.

The type of entrepreneur discussed in this paper is best described as a 'supply chain innovator' rather than an arbitrageur, because they initiate a linkage between customers and producers. They restructure demand and reallocate resources. They induce both substitution effects and complementarity effects. By innovating a novel product, an entrepreneur encourages customers switch away from the alternative products, i.e. they encourage substitution in demand. By placing orders for the novel product they induce producers to switch resources into the production of the novel product, i.e. they effect substitution in supply. These substitutions are complementary movements which release just sufficient resources into production to accommodate the shift in demand towards the novel product. This is a discrete analogue of a smooth and continuous 'neoclassical' substitution effect.

This discussion has emphasised the role of the entrepreneur in finding solutions to customer problems, but the analysis also applies, of course, to finding customers who may welcome novel products that exploit the potential of new technologies to generate new high-tech products or supply existing products at much reduced prices.

Comparison with Schumpeterian theory

Schumpeter distinguished five main types of innovation, all of which can be analysed using the type of model set out above. These include the discovery of new export opportunities and new sources of raw materials, as discussed in some of the cases presented above. Modern Schumpeterian theory has been mainly focused in technological innovation, however. In Schumpeterian theory the entire market for a novel product can potentially be pioneered by a single entrepreneur. As a result, segmentation is often ignored in modern Schumpeterian theory [De Vecchi and Stone, 1995]. However, recognition of segmentation is essential where the diffusion of innovations is concerned. It is because of segmentation that innovations diffuse slowly rather than instantaneously. An innovation may spread out like a wave-form, where each new replication stimulates further replication until the innovation has saturated the entire global market. The theory presented in this paper is well-adapted to analyse the diffusion of innovation in this way.

The theory can, indeed, do better than this. It is possible, for example, to introduce exogenous changes in the global environment which induce a series of changes in the global economy driven by entrepreneurial responses to new conditions. These exogenous changes may stimulate the innovation of new technologies and new products that exploit new opportunities or address emerging problems.

Each new generation of entrepreneurs can learn from previous generations. They can develop new products and technologies that make better use of scarce resources or better fulfil consumer requirements. They can also modify existing products and adapt existing technologies too. The applications discussed in Part Three demonstrate how models of segmented global markets can be applied to analyse the evolution of the global economy from one generation to the next.

The morality of the market

Segmentation theory also illuminates the debate on the morality of markets. Philosophical individualists often highlight the role of greed in stimulating consumer demand and the role of selfishness in bargaining successfully for a favourable price. This has led many economists to view the market process as an adversarial one, as noted earlier.

The theory presented above has been quite explicit that entrepreneurs are motivated by profit. This does not mean, however, that they are motivated by greed, or self-interest, as well. Indeed, it does not rule out the possibility that they are motivated by doing good. An entrepreneur could establish a business to make profit in order to finance a charity, e.g. there are many entrepreneurial charity shops on British high streets that maximise profit in order to fund charitable activities. An entrepreneur could also establish a business simply because they prefer to be 'independent ' and to be 'their own boss'; i.e. they prefer to own and control their own business rather work for someone else, simply as a 'life-style choice'.

It could, however, be argued that entrepreneurs who are motivated by greed or self-interest are most likely to create large businesses that are highly profitable. There are, in fact, a great many motives that could stimulate a successful entrepreneur, including the pursuit of power, celebrity, fame and fortune, and even a quest for immortality. There is however, an ambiguity, as to whether greed breeds success or whether success make people greedy for more.

The 'gains from trade' approach suggests a rather different perspective on these issues because it highlights the benefits of co-operation through exchange. Everyone potentially benefits if everyone can specialise according to their personal comparative advantage and consume according to their personal tastes. Markets therefore 'coordinate' or 'harmonise' the activities of different individuals. Within the gains from trade approach, the main moral problem is the distribution of the gains from such coordination. There are two aspects of this, relating to the process and to the outcome.

Adversarial bargaining on individualistic principles is a problem, because haggling uses up time and energy that could be devoted to more productive use. By intermediating the market, and stipulating the prices at which they will buy and sell, market-making entrepreneurs therefore provide a valuable service. They charge for this service by setting a margin between the wholesale and retail price. This margin is constrained by competition between them. Substituting competition between well-informed entrepreneurs for competition between ill-informed individual buyers and sellers therefore contributes to harmony in the market place.

The second issue is whether the outcome of this intermediation process is 'fair'. This is not a problem for the individualist, because fairness implies a social judgement, which is irrelevant if 'there is no such thing as society'. At the other extreme, socialists have condemned the entrepreneur on the spurious grounds that, as they do no manual work, they do not deserve a reward. The time they devote to information processing and decision-making, and the risks they bear, are simply ignored.

The theory of market segmentation shows that in any market the rewards appropriated by entrepreneurs depend crucially on the degree of competition between them [Alvares, 2007]. In a small local market there may be only a single entrepreneur, who therefore has a monopoly, while in a large urban market there may be much more competition. Competition may reduce profit margins, and so the distribution of income may therefore be 'fairer' in the latter case. In general, the smaller the market, the more scope there is for monopoly or collusion, and so the more unequal, and therefore 'unfair', the distribution of income is likely to become.

Overall, the theory of a segmented markets provides a nuanced account of market outcomes and therefore facilitates a detailed examination of the equity of income distribution. It is unnecessary to focus on the overall factor distribution of income between classes, such as capitalists and workers, which is characteristic of neoclassical economics. Instead, it is possible to analyse the distribution of income at a more disaggregated level, both within and between individual segments of the economy.

PART FIVE: EXTENSIONS OF THE ANALYSIS

The social context of entrepreneurship

The analysis so far has emphasised the 'hard' skills of the entrepreneurs, in terms of collecting and synthesising information, creating linkages, and setting prices. These hard skills are often complemented by soft skills, however. The theory presented in this paper does not have anything particularly original to say about the social context of entrepreneurship, but it does have something to say about the specific social skills that a successful entrepreneur is likely to require [Minniti and Bygrave, 1999; Bygrave and Minniti, 2000]. Examples of such skills are making contact with people who have access to confidential information, and making agreements with competitors to restrict competition.

Information networks and clubs

An entrepreneur will typically identify an opportunity through a synthesis of information. No single source of information is sufficient to identify an opportunity. At the very least an opportunity involves a latent, or unfulfilled demand, the ability to supply a product (e.g. contacts with producers) and the means of establishing a linkage between the two. There may be many independent, but complementary, sources of information that need to be accessed and analysed [Watts, 2003].

Access to information is always a potential problem, as key information of strategic value will often be withheld by those who possess it. Furthermore, the opportunity to analyse the information may be limited too. It requires expertise to decode specialised sources of information, and acquiring the relevant experience can be a time-consuming process. It is therefore useful to access an expert – someone who is familiar with the source and can

explain its significance. Experts are often in short supply and may be very much in demand, so gaining access to them may be difficult. In other words, the synthesis of information is a social process as well as an investigative one [Burt, 1992; Hidalgo, 2015].

To gain access to people who hold confidential information, or possess special skills, it is often useful to join a club. A club is a special type of institution which facilitates interaction between people with similar interests. In some cases access to a club may be free, in some cases there may be a fee, and sometimes election to membership may be required.

There is a well-established body of economic theory on clubs [Buchanan, 1968; Olson, 1965]. A club is typically an association with fee-paying members which provides services or facilities that are shared between the members. The establishment of a club may itself be regarded as an entrepreneurial activity. Like entrepreneurs, clubs effect coordination, because they make their members better off, without making other people worse off in the process. If this were not the case then members would not join and pay their membership fees.

Clubs may compete against each other in the same way as firms, whether or not they are privately-owned. However, an important distinction between firms and clubs is that, not only are the owners in contact with the members, but the members are in contact with each other. Members of a firm may be in contact with others within their team, but they may not be in contact with people from different teams. Because the members of a club can socialise, the owner or president of a club can be readily held to account; indeed, in a serious situation members can discuss their grievances with each other and protest as a group. By contrast, employees within a firm may be locked into small teams that have little influence because their members have no opportunity to socialize with those outside their team.

The role of the entrepreneur in establishing a club

The establishment of a club may be considered entrepreneurial, whether the club operates as a business or a charity. The entrepreneur must identify a group of people with common interests (social, political, professional, business, or leisure). They may contact potential members personally, and invite new members to 'tell their friends'; alternatively they may advertise impersonally.

When considering whether to establish a club an entrepreneur must assess how many people will be willing to join, the membership fees they will be willing to pay, and what services they will expect in return. They must also assess how much these services will cost to provide. Their motives may be selfish or altruistic; in either case they will need to assess the probable impact of the club on the potential beneficiaries.

There may be competition between clubs in the same area of interest, e.g. local 'chambers of commerce' may compete for members in nearby towns, while rival professional associations may compete to regulate their profession and advance their members' interests. The analysis of competition presented earlier is therefore also relevant to clubs.

Clubs act as hubs in social networks. While many local clubs support isolated or lonely people, the hubs that are most important to entrepreneurs are those where they can meet with wealthy and well-informed people. Access to privileged information is always useful, but for entrepreneurs it is particularly valuable, as they are well-equipped to exploit such information for business purposes. One of the great advantages of living in or near a great metropolis is

that it is an information hub and many of the wealthiest and most knowledgeable individuals will tend to meet up there. Access to such people not only provides inside information and access to capital, but also an opportunity to lobby for changes to law or government policy which favour an entrepreneur's activities [Aldridge, 1987].

The role of clubs in promoting coordination is illustrated in Figure 7. The figure relates to one of the scenarios discussed above, in which the entrepreneurs $E1$, $E2$ compete for supplies from producers $B1$, $B2$, in order to meet the demands of the two customer groups, $A1$ and $A2$, which they each monopolise.

FIGURE 7 HERE

The figure shows that $E1$ has just joined the club $C1$ in order to gain understanding of the problems encountered by members of $A1$. They have chosen $C1$ because they believe that people who have the kind of problem that needs to be solved are likely to join this particular club. For example, a publisher of novels may join a literary club, while a supplier of outerwear may join a hiking club. In practice, of course, it is not always necessary, or appropriate, to join a club because the information required can be obtained in other ways. An entrepreneur who plans to launch a new alcoholic drink, for example, can simply visit bars and public houses and chat informally to their patrons.

While $E1$ is researching $A1$ by joining $C1$, $E2$, unaware of $E1$, and unknown to them, is doing similar research by joining club $C2$. When $E1$ and $E2$ have determined the key attributes that their product must possess they need to make contact with suitable producers, $B1$ and $B2$. There are two potential clubs at which they could meet potential producers, namely $D1$ and $D2$. Both potential producers, $B1$ and $B2$, are members of both clubs, it is assumed, so that membership of either club will do. In this case each entrepreneur is likely to join their local club; thus $E1$ will join $D1$ and $E2$ will join $D2$.

Time lags in production and distribution may mean that receipts from sales lag behind payments to producers, so that working capital tied up in inventory needs to be financed. In addition, there may be set-up costs incurred on entry, although these were omitted from the models presented above. Each entrepreneur may therefore gain by joining a club, such as $F1$ or $F2$, where introductions can be made to potential lenders and investors. As a result, $E1$ is may make contact with a member of $F1$ and, likewise, $E2$ with a member of $F2$.

The membership of $F1$ and $F2$ may include some well-connected people who may not only provide finance but also give access to the government or an industry regulator. Access to government is particularly important in sectors such as construction and transport, where government finances many projects out of taxation. If $E1$ and $E2$ are located in different countries then each will normally get in touch with a different national government, represented by $G1$ and $G2$ respectively. The controlling political party may, in some respects, be similar to a club, and it may be more rewarding for an entrepreneur to have contacts in a governing political party than with officials in a department of government. Some members of the government $G1$ may belong to $F1$ and some members of $G2$ may belong to $F2$.

This analysis of clubs establishes a link between the theory of entrepreneurship and the theory of social networks. This link was well understood some fifty years ago, and was reflected in the literature on mathematical sociology and organizational studies [Coleman, 1964; Trenberth, L, 2009].

It has been shown that clubs help to create entrepreneurs and that entrepreneurs create clubs. To maintain an inter-generational equilibrium it is necessary that each generation of clubs creates the next generation of entrepreneurs and that each generation of entrepreneurs maintains sufficient clubs. This does not require that clubs persist indefinitely but only that new clubs are regularly established to replace those clubs that close down. Indeed, some degree of turnover in clubs is desirable, as the composition of the population of clubs may need to change as the types of opportunities that arise and the types of problems that confront entrepreneurs change over time.

Collusion

In the models presented above an entrepreneur who is the first to enter a market can, in principle, appropriate all the gains from trade through their monopoly and monopsony power. With two competing entrepreneurs and excess supply, however, one of the entrepreneurs will make no profit and the other will normally earn less profit than they would have done as a monopolist. Both entrepreneurs could potentially benefit if they agreed to limit, or even eliminate, competition between them. They could then purchase at the producer's unit cost and sell at the customers' reservation price. Conflict would remain, however, over how to share the profit between them.

This raises the issue of bargaining skills. The entrepreneurs may agree to 'split the difference' and take half the overall profit, but in practice the first-mover may attempt to bargain for a larger share. There is scope, however, for each of the entrepreneurs to cheat the other, e.g. by under-stating their actual revenues or overstating their actual costs. This in turn raises the issue of trust.

Trust is an important factor in entrepreneurship. Winning the trust of customers is important, although a deceitful entrepreneur may be able to win the trust of credulous customers. Likewise winning the trust of producers (e.g. by paying promptly) is important too. But the crucial factor where collusion is concerned is winning the trust of other entrepreneurs. An entrepreneur who is known to deceive their customers or their producers may find it difficult to win the trust of other entrepreneurs with whom they wish to collude [Nooteboom, 2002; Herold, 2019]. Since the members of a secret cartel set out to deceive other people there is a strong suspicion that they may be willing to deceive each other too.

Organizing and financing collusion

There are different ways organizing collusion. They range from informal 'understandings', through cartels and trusts to joint ventures and mergers.

Cartels may be secretive, or very public government-sponsored organizations, or positioned somewhere in between. Some cartels fix members' prices, and some fix quantities, while others allocate monopolies of specific segments to specific entrepreneurs (e.g. on a national or regional basis). Some cartels have a formal bureaucracy that monitors members' behaviour, while others may rely on informal 'gentlemen's agreements' [Buckley and Casson, 2021; Liefmann, 1932; Stocking and Watkins, 1946].

Trusts are organizations, often owned by banks or wealthy individuals, that hold significant financial stakes in different businesses run by different entrepreneurs, and use their influence

(often exercised through directors on an executive board) to coordinate the strategies of competing firms.

These approaches to collusion are largely consensual, but similar results can be obtained through hostile take-overs of competing firms. Hostile take-overs often require substantial loans, made to the acquiring firm, to finance the purchase of voting shares in the target firm. These loans are not unlike the loans that are made to finance the expansion of a high-growth small firm. In each case the financier must assess the entrepreneurial skills of the people controlling the enterprise they propose to finance. Thus while a conventional entrepreneur assesses an opportunity to enter a market, a financier must assess the opportunity to finance a take-over. A key element in this is to assess the prospects of the target firm, and the ability of the acquiring entrepreneur(s) to manage a larger business than they managed before.

PART SIX: SUMMARY AND CONCLUSIONS

A metaphor

There is a nineteenth-century metaphor, whose origin is obscure, which likens the entrepreneur to the builder of a watch tower, who then ascends the watchtower to get a better view of their surroundings. From their superior vantage point they notice that down on the ground things do not work as smoothly and efficiently as they should. The people on the ground cannot recognise this because their horizons are much more limited and their view is often obstructed. The entrepreneur identifies numerous opportunities for improvement, but has to prioritise them as they cannot exploit them all at once.

These opportunities may involve people making swaps (participating in exchange and trade) or changing jobs to produce things that are in short supply (specialisation in the labour market). They may involve developing new products, or producing existing products using cheaper methods. But someone has to initiate these activities, because the individuals on the ground cannot recognise these opportunities for themselves. The people who need to make contact are not necessarily neighbours; they may be distributed across the entire range of vision. To bring these people together the entrepreneur needs to set up shop at a central place, which will normally be close to the foot of their watchtower. This central place will act as a communications hub where information from the surrounding areas is concentrated. This allows the entrepreneur to make the synthesis of information from which opportunities for coordination can be identified [O'Driscoll, 1977]. As their business expands the entrepreneur may set up satellite hubs as well.

The models presented in this book are also metaphors. The language, however, is very different. It is not a picture that has been painted, but a mathematical model that has been presented. The watchtower is built on firm foundations, and likewise the mathematical model is firmly grounded in the laws of logic and arithmetic.

Key results

This paper has attempted to address key issues in the theory of entrepreneurship through a direct confrontation with mainstream neoclassical economic theory. The focus has been on the 'market-making' entrepreneur. The main results may be summarized as follows.

Most markets operate through two prices rather than one: a retail price and a wholesale price. Mainstream economics, by contrast, normally assumes a single uniform price.

Supply chains are key. Many products are brought to market through a chain which comprises both upstream producers and downstream wholesalers and retailers, with each link in the chain being mediated by an entrepreneur. Each chain corresponds to a 'route' or 'linkage' in the mathematical model set out in Part Two.

Markets are segmented in all sorts of different ways. In some cases customers simply do not know the price at which a product is being sold in a different segment, and in other cases they may know that it is cheaper but realise that they cannot access it. They might also know that it is dearer, but if they face transaction costs and transport costs there may be no profit in arbitrage. In mainstream theory, by contrast, market segmentation is mostly ignored. There is a theory of retail location which is related to spatial segmentation, but this is typically formulated in terms of urban economics rather than as a theory of market segmentation as a whole. The only coherent account of market segmentation in mainstream economics is available from the international economics literature, but this focuses almost exclusively on segmentation created by international transport costs and tariffs.

Markets exhibit price dispersion. Price dispersion is a consequence of market segmentation. Mainstream economics explains price dispersion using consumer search theory. Search theory introduces price dispersion 'by the back door.' It is treated as an awkward fact of life that needs to be explained, rather than as a fundamental characteristic of segmented markets. Because there are no entrepreneurs, mainstream theory assumes that customers are reactive to the behaviour of sellers, rather than the other way round [Bliss, 1988; Bresnahan and Reiss, 1985, Burden and Judd, 1983, Salop and Stiglitz, 1977].

There is no need to model duopoly or oligopoly using conjectural variations. In the model presented in this paper price and quantity are set by a very different mechanisms to those assumed in conventional duopoly and oligopoly theory. When an entrepreneur is negotiating with customers they know each customer's reservation price (which is the same for all the customers in a given segment) and the total demand in their segment. They also know from experience whether there is a rival entrepreneur bidding to serve the same market. By contrast, the customers do not know the entrepreneur's cost of procurement, and cannot collude against them because they are not in direct contact with producers. Provided that the model specifies exactly how an entrepreneur responds to the information signals they receive, price and quantity are determinate. The outcome differs from a conjectural equilibrium, in which rival sellers are aware of each other's identity and can correctly predict how their rivals will respond to any price or quantity to which they commit [Bertrand, 1883; Cournot, 1897; Vives, 1999].

Marketing and distribution are entrepreneurial functions which deserve explicit recognition. This paper has focused on a particular type of entrepreneur, namely the market-making entrepreneur. Their particular skill lies in identifying demand for novel products, or new sources of demand for existing products. Such activities require resources such as advertising, display, order-taking and delivery, collection of payment, etc. all of which use up resources. The procurement of these resources incurs access costs. The strategic advantages possessed by an entrepreneur include not only the ability to identify new sources of demand and supply, but also the ability to make contact with those concerned, negotiate with them and win their trust. These activities are often regarded as mundane, but in practice they can be crucial in gaining a 'competitive edge' over rival entrepreneurs.

Markets may be internalised; e.g. the same entrepreneur may own and control successive stages of production. In this context, internalisation of distribution exemplifies forward integration, while internalisation of production exemplifies backward integration. An entrepreneur may also integrate into research and development. Other types of contractual arrangement can be adopted too. For example, an entrepreneur may adopt a 'putting out' system, involving serial subcontracting, in which successive stages of production are delegated to different independent specialised producers. The entrepreneur retains ownership of the semi-processed product, but labour and machinery are provided by the subcontractors at their own risk. Mainstream economics, by contrast, has relatively little to say about internalization and often implicitly assumes that the entrepreneur and the producer are one and the same person.

There is no need to model human behaviour on utilitarian principles. In an innovative economy a person's utility function will either have to keep changing, or to be so general that it encompasses every imaginable product whether it is available or not. By basing the analysis of consumer demand on utility maximisation, mainstream economics has placed a nebulous psychological concept at the heart of its theory, and in doing so it has cut itself off from other social sciences. The utility function may be replaced by a set of mutually-consistent decision rules, as exemplified in the case studies presented in this paper. Mutual consistency is sufficient to preserve the 'rationality' of human behaviour that is a characteristic assumption of neoclassical economics. Irrationality may then be explained, as appropriate, by lack of information, incorrect information, or a failure to interpret information appropriately.

If utility were interpersonally comparable then it would be possible to use the maximization of total utility as a criterion of social efficiency, but since it is now agreed that it is not, it is necessary to find some other criterion of social efficiency instead. Assuming that everyone is a clone of the same representative individual, as in some branches of neoclassical theory, is not an adequate basis for an analysis of market segmentation either. This paper has proposed that maximising the gains from trade, as measured by the increase in the total value of consumption derived from a given set of resources, is a superior alternative, which has already been successfully applied to theories of international trade. A theory of entrepreneurship can then be based on the principle that an efficient economy will maximise the gains from trade conditional on the opportunities for trade that have been identified by entrepreneurs.

The emergence of market structure in a competitive product market can be modelled using linear programming. It has been shown that when customers have rectangular demand schedules and producers have rectangular supply schedules, and access by entrepreneurs to customers or producers incurs constant unit costs, the market outcome can be determined by maximising the total gains from trade subject to a set of linear constraints. These constraints are determined by satiation demands in each customer segment, the capacity of each producer, and the availability of linkages. The parameters that appear in the constraints are the reservation prices of each customer group, the unit costs of each producer and the unit access costs of each linkage. The solution of the programming problem indicates which of the available linkages are utilised and which are not, and how much product flows along each of them. It also indicates the gains from trade that are generated by each entrepreneur. Linear programming models have previously been used mainly to optimise the choice of production

techniques within a firm, which has led many scholars to assume, quite wrongly, that the method cannot be applied to trade. It has been known for many years, however, that linear programming can be applied to models of international trade, although previous applications were mainly confined to trade in intermediate products rather than trade in final products, as analysed here.

The prices and profits generated by the market can be determined from the principles of competitive price adjustment. Where there is potential competition a successful entrepreneur must bid down the customer price, and/or bid up the producer price, so that no other entrepreneur can afford to undercut them in the retail market or outbid them in the wholesale market. Where entrepreneurs are obliged to adjust their prices in this way, the potential profit they have lost will accrue to the customers and/or producers instead. In other words, the greater the strength of competition the smaller the profit accruing to entrepreneurs and the greater the surplus accruing to customers and producers instead. Another way of expressing this point is to say that the profit accruing to an entrepreneur stems from the superiority of the linkages they control over the best competing linkages operated by other entrepreneurs. Their profit is therefore a measure of their 'competitive advantage' over their closest rivals. If, in the absence of the entrepreneur, the segment(s) concerned would not be served at all, the entrepreneur's reward may be said to represent a 'monopolistic advantage' instead. In existing entrepreneurship literature this distinction between competitive advantage and monopolistic advantage is not always clear.

The evolution of markets can be modelled as a sequence of entries (*and possibly withdrawals*) *by successive entrepreneurs.* In some cases incumbent entrepreneurs may fail to respond to subsequent entry, but it is more realistic to assume that incumbents have an opportunity to react to each entry after it occurs. If an incumbent is unable to adjust, and thereby loses all its market, it may decide to remain dormant in case new opportunities arise, or it may decide to quit. This decision may depend on whether the entrepreneur has entry costs, or other costs, that they could recover if they quit. In some cases entrepreneurs may be able to anticipate the responses of rivals before they commit resources irreversibly to some specific use, but in other cases they cannot.

Because the theory of entrepreneurship involves many parameters, the theory can be finely tuned to address a wide range of specific cases. In contrast to mainstream economic theory, where models typically involve restrictive assumptions and very few parameters, entrepreneurship theory involves relatively few restrictive assumptions and many more parameters, which means that applications can be tailored to many individual real-world situations. This also makes the theory easy to teach, because although it involves the use of mathematical methods, the application of these methods can be illustrated using engaging practical and topical applications. In other words, it offers a rigorous alternative to the traditional case study approach.

Implications for research

In management studies and the social sciences today it is quite usual to develop and present theory in purely conceptual terms. A problem is solved by introducing a new concept or creating a new buzz-word [Casson and Della Giusta, 2011]. In reality, however, only a rigorous theory can guarantee a logically consistent solution to a complex real-world problem. Few entrepreneurship scholars have used abstract mathematical methods to develop

their theories. The absence of mathematical proofs renders entrepreneurship theory vulnerable to the criticism that it rests on plausible arguments rather than on a formal logical structure. The lack of formal logical structure can be cited by critics as evidence that the theoretical foundations of the theory are weak.

This paper has presented a mathematical model – indeed, a family of mathematical models – which provide explicit logical foundations for an economic theory of entrepreneurship. This theory allows entrepreneurship scholars to use tools and techniques that are in many respects superior to those employed by mainstream economists. Instead of appealing to the 'invisible hand' of Adam Smith, the fictional auctioneer of Walrasian economics, or the 'tendencies' of a Marshallian market, the theory focuses on entrepreneurial rivalry in segmented markets. Segmentation is a 'fact of life' in many markets and is largely ignored by mainstream economists, except in respect of international and inter-regional trade. Segmentation is not the same as product differentiation, however. Although product differentiation may be considered to be a special case of segmentation, it is usually modelled by assuming that each variety has a similar demand, which is a highly restrictive assumption that is avoided by segmentation theory.

The theory presented here allows for different entrepreneurs to have different personal capabilities. Because all entrepreneurs are different, it is possible to introduce personal characteristics into the theory. Entrepreneurs differ in personality, family background, cultural heritage, and the range of social contacts they possess. As a result, some entrepreneurs are more aware of the market environment than others, and can therefore make contact with a wider range of customer groups or a greater number of producers. Similarly, some entrepreneurs are better equipped to cooperate or collude with others through market-sharing agreements or other restrictive business practices, and have better access to social networks through which they can raise finance.

By drawing a clear distinction between the entrepreneur – a person – and the firm – an institution – the strategic advantages of vertical and horizontal integration are highlighted. The scope of the firm is determined not only by the visionary capabilities of the entrepreneur but also by the challenges that the entrepreneur faces in operating linkages – both forward, backward and horizontally. Different entrepreneurs may adopt different degrees of 'internalization' or 'integration' because of the different capabilities – both entrepreneurial and managerial – that they possess.

Entrepreneurship theory therefore has a rightful claim to provide best available account of market behaviour, but, as emphasised above, it cannot prove that claim unless it can provide a formal model that addresses all of the key questions concerning that topic. It was noted at the outset that mathematicians and abstract theorists have a legitimate point when they say that ordinary words have multiple connotations, which makes it easy for people who rely exclusively on everyday language to string together seemingly plausible sentences that on close examination are either ambiguous, or over-simplify a complicated situation.

Entrepreneurship theorists sometimes write as if they possess profound philosophical insights into human nature (in particular, entrepreneurial motivation) and human society (in particular, political freedoms and legal institutions), but these claims ring hollow if they cannot even provide a coherent account of how market prices are determined using a mathematical model. This suggests that future progress in the development of entrepreneurship theory is most

likely to be achieved by the refining the mathematical approach presented in this paper, rather than in perpetuating ideological debate over personal ambition, freedom of choice, the rule of law and the virtues of deregulated capitalism.

Implications for teaching

The models presented in the paper, and the methods used to solve them, suggest that it is no longer necessary to teach economics and entrepreneurship as separate subjects offering rival accounts of firms and markets. The theory of entrepreneurship summarized in this paper includes all the key elements of core economics: scarcity, choice, coordination, specialisation, comparative advantage, division of labour, gains from trade, optimisation, profit and Pareto efficiency. The two key elements of a first-year economics course that it does not include – namely the utility function and the Cobb-Douglas production function – are redundant. The utility function is replaced by explicit demand parameters, while the Cobb-Douglas production function is replaced by a set of specific well-defined technologies.

The theory described above shows that Marshall's description of market equilibrium as a balance between two opposing forces of demand and supply is in some respects misleading [Marshall, 1890]. Markets facilitate the pursuit of mutual benefit, and efficient markets maximise the gains from trade. Equilibrium is achieved when no further profit opportunity remains to be exploited. Marshall's focus on a uniform market price, which stems from an assumption that information is a common good within a market place, is gratuitous. A global market can be in equilibrium when prices are dispersed so long as information is incomplete. The equilibrium condition is that no further trade is profitable along any linkages of which anyone is aware.

The so-called 'harmony' of the market outcome is therefore achieved, both in theory and in practice, not by impersonal forces, but by entrepreneurs. There is, however, potential for conflict over the distribution of the gains from trade. If a market is dominated by a single entrepreneur then they can, in principle, appropriate all the gains as profit. Competition from rival entrepreneurs places a check on this, however. Incumbent entrepreneurs can, though, attempt to erect barriers to entry in order to maintain their market power. In some cases society, acting through government and the law, decides to condone this practice, in order to incentivize innovation and risk-taking, while in other cases it discourages it, e.g. because it reduces efficiency and exacerbates inequalities in the distribution of income.

It could be objected, nevertheless, that this economic approach to entrepreneurship remains too mathematical to provide analytical foundations for entrepreneurship theory. However, the examples given above illustrate that the theory can be expounded using a variety of practical examples. This demonstrates the power and generalisability of the theory, which allows researchers to adapt the formal model to a wide variety of different cases in different fields of study. One reason why the theory is so flexible is that it contains many parameters, so that adjusting the range of these parameters allows the model to applied to many different situations. Another reason is the theory's basic simplicity: namely the fundamental principle of maximising the gains from trade subject to constraints imposed by segmentation, which reflects both physical obstacles to trade and the partial and unequal distribution of information within society.

Even more significant, perhaps, is the fact that neither the formulation nor the application of the application requires the use of calculus. This is because of the linearity of the model. Despite the apparent abstraction of an *N*-dimensional convex set, and a tangent hyperplane, the basic model can be illustrated in two or three dimensions if required. This version of the model is very simple, however, and although a two-dimensional projection of a three-dimensional convex set is difficult to draw, the exercise can aid the imagination and enhance intuitive understanding of the way in which the model works.

Some readers may go so far as to object to the use of algebra in the formulation of the model, but even this can be avoided to some extent. The key points can all be made using practical numerical examples, as illustrated in Part Three. If necessary the solutions can be programmed in Excel on a laptop; indeed, it is only a matter of time before solutions are available using a mobile phone.

Conclusions

Entrepreneurship theory addresses problems concerning prices and quantities, revenues and costs, profits and losses, and the appropriation of the gains from trade in general. The significance of such numerical measures reinforces the need for, not only logical argument, but explicit mathematical theory. This theory needs to be grounded in the realities of everyday life, however. The world of modern neoclassical economics is an artificial world dominated by assumptions that are made for mathematical convenience rather than practical relevance. This paper has presented an alternative type of model that is just as rigorous but grounded in reality [Morgan, 2012; Weintraub, 2002].

The advantages of the theory set out in this paper are that

- it sufficiently flexible to be adaptable to a wide range of real-world situations;
- it can analyse not just a steady state but an evolving sequence of states;
- it determines the value of a wide range of economic variables;
- it generates new types of information, such as price dispersion and the personal distribution of income; and
- it employs a model that can be solved using simple techniques, including one available on a laptop computer.

Because it integrates entrepreneurship and economic theory the theory can be used in both business schools and economics departments Because it is so versatile it is well-adapted to case study analysis. Because it can analyse sequential interactions it is well adapted to teaching and researching business strategy. Overall, it reconstructs significant parts of mainstream economic theory to reveal entrepreneurship at the heart of the market economy.

REFERENCES

Alchian, A.A. and H. Demsetz (1972) Production, information costs and economic organization, *American Economic Review*, 62 (5), 777-795

Aldridge, H.E. (1987) The impact of social networks on business founding and profit: A longtitudinal approach, in N.W. Churchill et al (eds.) *Frontiers of Entrepreneurship Research*, Wellesley, MA: Babson College, 1454-168

Alvarez, S.A. (2007) Entrepreneurial rents and thr theory of the firm. *Journal of Business Venturing*, 22(3), 427-442

Arthur, W.B. (1989) Competing technologies, increasing returns and lock-in by historical events, *Economic Journal*, 99, 116-131

Backhaus, J.C. (2003) *Joseph Alois Schumpeter: Entrepreneurship, Style and Vision*, Boston: Kluwer

Bacon, R.W. (1984) *Consumer Spatial Behaviour*: *A Model of Purchasing Decisions over Space and Time*, Oxford: Oxford University Press

Baligh, H.H. and L.E. Richartz (1967) *Vertical Market Structures*, Boston: Allyn & Bacon

Baumol, W.J. (1968) Entrepreneurship in economic theory, *American Economic Review*: *Papers and Proceedings*, 58, 64-71

Baumol, W.J. (1990) Entrepreneurship: productive, unproductive and destructive, *Journal of Political Economy*, 98, 893-921

Baumol, W.J. (1994) *Entrepreneurship, Management and the Structure of Payoffs*, Cambridge, Mass: MIT Press

Bertrand, J. (1883) Book review of 'Theorie Mathematique de las richesse Sociale', *Journal de Savants*, 67, 499-508

Bhide, A.V. (2000) *The Origin and Evolution of New Businesses*, Oxford: Oxford University Press

Bliss, C. (1988) A theory of retail pricing, *Journal of Industrial Economics*, 36 (4), 375-391

Bloch, F. and H. Ferrer (2001) Trade fragmentation and coordination in strategic market games, *Journal of Economic Theory*, 101, 301-316

Boulding, K.E. (1960) *Linear Programming and the Theory of the Firm*, New York: Macmillan

Bourdieu, P. (1977) *Outline of a Theory of Practice*, Cambridge: Cambridge University Press

Bowles, S. and W. Carlin (2020) What students learn in Economics 101: Time for a change, *Journal of Economic Literature*, 58(1), 176-214

Bresnahan, T.F. and P.C. Reiss (1985) Dealer and manufacturer margins, *Rand Journal of Economics*, 16 (2), 253-268

Brueghel, P., M. Andersson and J. Eklund (2022) Pioneering entrepreneurship research: How, by whom and when, *Foundations and Trends in Entrepreneurship*, 18 (2), 75-158

Buchanan. J.M. (1968) *The Demand and Supply of Public Goods*, Chicago: Rand McNally

Buckley, P.J. and M. Casson (1976) *The Future of the Multinational Enterprise*, London: Macmillan

Buckley, P.J. and M. Casson (2021) Multinational enterprises and international cartels: The strategic implications of de-globalization, *Management and Organization Review*, 167 (5), 968-988

Burdett, K. and K.L. Judd (1983) Equilibrium price dispersion, *Econometrica*, 51, 955-969

Burt, R.S. (1992) *Structural Holes*: *The Social Structure of Competition*, Cambridge, MA: Harvard University Press

Bygrave, W.D. and M. Minniti (2000) The social dynamics of entrepreneurship, *Entrepreneurship Theory and Practice*, 24 (3), 25-36

Casson, M. (1982) *The Entrepreneur: An Economic Theory*, Oxford: Martin Robertson, (2nd. ed., 2003, Cheltenham: Edward Elgar)

Casson, M. (1997) *Information and Organization*: *A New Perspective on the Theory of the Firm*, Oxford: Oxford University Press

Casson, M. (2003) Marshall and Marketing, in J. Creedy (ed.), *From Classical Economics to the Theory of the Firm*: *Essays in Honour of D.P. O'Brien*, Cheltenham: Edward Elgar, 194-219

Casson, M. (2005) Entrepreneurship and the Theory of the Firm, *Journal of Economic Behaviour and Organization*, 58, 327-348

Casson, M. (2010) *Entrepreneurship: Theory, Networks, History*, Cheltenham: Edward Elgar

Casson, M. (2024) *Entrepreneurship and the Market Economy*: *A New Perspective*, Cheltenham: Edward Elgar

Casson, M. and M. Della Giusta (2011) Buzzwords in international management education, in Tang, D., H. Kazeroony and G. Ellis (eds.) *Routledge Companion to International Management Education*, Abingdon: Routledge, 344-357

Chamberlin, E. (1933) *Theory of Monopolistic Competition*, Cambridge, MA: Harvard University Press

Clark, J.M. (1940) Towards a concept of workable competition, *American Economic Review*, 30 (2), 241-256

Clark, J.M. (1961) *Competition as a Dynamic Process*, Washington, DC: Brookings Institution

Coase, R.H. (1937) The nature of the firm, *Economica*, (New series), 4, 386-405

Coleman, J.S. (1964) *Introduction to Mathematical Sociology*, New York: Free Press

Commons, J.R. (1924) *The Legal Foundations of Capitalism*, New Work: Macmillan

Cournot, A.A. (1897) *Researches into the Mathematical Principles of the Theory of Wealth* (trans, N.T. Bacon, ed. I. Fisher), New York: Macmillan

Creedy, J. (1986) *Edgeworth and the Development of Neoclassical Economics*, Oxford: Blackwell

Creedy, J. (1992) *Demand and Exchange in Economic Analysis*: *A History from Cournot to Marshall*, Aldershot: Edward Elgar

Creedy, J. (1998) *Development of the Theory of Exchange*, Cheltenham: Edward Elgar

Cusumano, M.A., A. Gower and D.B. Yoffie (2019) *The Business of Platforms*, New York: Harper Business

Dantzig, G.B. (1991) Linear programming, in J.K. Lenstra, A.H.G. Rinnooy and A. Schrijver (eds.) *History of Mathematical Programming*: *A Collection of Personal Reminiscences*, Amsterdam: Elsevier Science

De Vecchi, N. and A.J. Stone (1995) *Entrepreneurs, Institutions and Economic Change: The Economic Thought of J.A. Schumpeter* (1905-1925), Aldershot: Edward Elgar

Dorfman, R. (1951) *Application of Linear Programming to the Theory of the Firm*, Berkeley: University of California Press

Dorfman, R., P.A. Samuelson and R.M. Solow (1958) *Linear Programming and Economic Analysis*, New York: McGraw-Hill

Edgeworth, F. Y. (1881) *Mathematical Psychics*, London: C.K. Paul

Edgeworth, F.Y. (1925) *Papers relating to Political Economy*, London: Macmillan for the Royal Economic Society

Faias, M., C. Herves-Beloso and E. Moreno-Garcia (2010) Equilibrium price formation in markets with differentially informed agents, *Economic Theory*, 48, 205-218

Foss, N.J. and P.G. Klein (eds.) (2002) *Entrepreneurship and the Theory of the Firm*: *Austrian Perspectives on Economic Organization*, Cheltenham: Edward Elgar

Foss, N.J. and P.G. Klein (eds.) (2012) *Organizing Entrepreneurial Judgment*, Cambridge: Cambridge University Press

Foss, N., and P. Klein (2020) Entrepreneurial opportunities: Who needs them?, *Academy of Management Perspectives* 34, 366-377

Godley, A.C. and M. Casson (2016) 'Doctor, doctor …' entrepreneurial diagnosis and market-making, *Journal of Institutional Economics*, 11 (3), 601-621

Gould, J.P. (1980) The economics of markets: A simple model of the market-making process, *Journal of Business*, 53(3, Pt 2), S167-S187

Hamilton, S. and T. Richards (2009) Product differentiation, store differentiation and assortment depth, *Management Science*, 55 (8), 1368-1376

Harper, D.A. (1996) *Entrepreneurship and the Market Process: An Enquiry into the Growth of Knowledge*, London: Routledge

Hayek, F. A. von (1945) The use of knowledge in society, *American Economic Review*, 35, 519-530

Hebert, R.F. and A.N. Link (1982) *The Entrepreneur*: *Mainstream Views and Radical Critiques*, New York: Praeger

Herold, P. (2019) *Trust, Control and the Economics of Governance*, Abingdon: Routledge

Hicks, J.R. (1932) *Theory of Wages*, London: Macmillan

Hidalgo, C. (2015) *Why Information Grows*: *The Evolution of Order from Atoms to Economies*, New York: Basic Books

Hodgson, G. (2015) *Conceptualising Capitalism*: *Institutions, Evolution, Future*, Chicago: University of Chicago Press

Hotelling, H. (1929) Stability in competition, *Economic Journal*, 39 (153), 41-57

Ingrao, B. and G. Israel (1990) *The Invisible Hand: Economic Equilibrium in the History of Science*, Cambridge, MA: MIT Press

Kirzner, I.M. (1973) *Competition and Entrepreneurship*, Chicago: University of Chicago Press

Kirzner, I.M. (1979) *Perception, Opportunity and Profit*: *Studies in the Theory of Entrepreneurship*, Chicago: University of Chicago Press

Klein, P.G. (2010) *The Capitalist and the Entrepreneur*: *Essays on Organizations and Markets*, Auburn, Alabama: Ludwig von Mises Institute

Knight, F.H. (1921) *Risk Uncertainty and Profit*, Boston: Houghton Mifflin

Koopmans, T.J. (ed.) (1951) *Activity Analysis of Production and Allocation*, New York: Wiley

Krugman, P. (1980) Scale economies, product differentiation and the pattern of trade, *American Economic Review*, 70 (5), 950-959

Langlois, R.N. and P. Robertson (1995) *Firms, Markets and Economic Change*: *A Dynamic Theory of Business Institutions*, London: Routledge

Liefmann, R. (1932) *Cartels, Concerns and Trusts*, London: Methuen

Marshall, A. (1890) *Principles of Economics*, London: Macmillan

Meade, J.E. (1952) *A Geometry of International Trade*, London: Allen & Unwin

Mill, J.S. (1848) *Principles of Political Economy*, London: J.W. Parker

Minniti, M. and Bygrave, W.D. (1999) The microfoundations of entrepreneurship, *Entrepreneurship Theory and Practice*, 23 (4), 41-52

Morgan, M.S. (2012) *The World in a Model: How Economists Work and Think*, Cambridge: Cambridge University Press

Nooteboom, B. (2002) *Trust*: *Forms, Foundations, Functions, Failures and Figures*, Cheltenham: Edward Elgar

Nooteboom, B. (2014) *How Markets Work and Fail*, *and What to Make of Them*, Cheltenham: Edward Elgar

North, D.C. (1991) *Institutions, Institutional Change and Economic Development*, Cambridge: Cambridge University Press

O'Driscoll, G. (1977) *Economics as a Coordination Problem*: *The Contributions of Friedrich Hayek*, Kansas City: Sheed Andrews McNeel

Olson, M., Jr. (1965) *The Logic of Collective Action: Public Goods and the Theory of Groups*, Cambridge, MA: Harvard University Press

Penrose, E.T. (1959) *A Theory of the Growth of the Firm*, Oxford: Blackwell

Porter, G. and H.C. Livesay (1971) *Merchants and Manufacturers: Studies in the Changing Structure of Nineteenth-century Marketing*, Baltimore: Johns Hopkins University Press

Ramoglou, S (2021) Why do disequilibria exist? An ontological study of Kirznerian economics, *Cambridge Journal of Economics*. 45 (4), 833-856

Ramoglou, S. and E.W.K. Tsang (2018) Opportunities lie in the demand side: Transcending the discovery-creation debate, *Academy of Management Review*. 43 (4), 815-818

Ricardo, D. (1817) *On the Principles of Political Economy and Taxation*, London: John Murray

Roberts, B. and B.R. Holden (1972) *Theory of Social Process*: *An Economic Analysis*, Ames, IA: Iowa State University Press

Salop, S. and J. Stiglitz (1977) Bargains and rip-offs: A model of monopolistically competitive price dispersion, *Review of Economic Studies*, 44, 493-510

Schumpeter, J.A, (1934) *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle* (ed. R. Opie) Cambridge. MA: Harvard University Press

Schumpeter, J.A. (1939) *Business Cycles*, New York: Wiley

Shackle, G.L.S. (1979) *Imagination and the Nature of Choice*, Edinburgh: Edinburgh University Press

Shane, S. (2003) *The General Theory of Entrepreneurship*: *The Individual – Opportunity Nexus*, Cheltenham: Edward Elgar

Shapley, L.S. and M. Shubik (1969) On market games, *Journal of Economic Theory*, 1, 9 -25

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

Stigler, G.J. (1942) *The Theory of Competitive Price*, New York: Macmillan

Stigler, G.J. (1957) Perfect competition, historically considered, *Journal of Political Economy*, 65(1), 1 -17

Stigler, G.J. (1961) Economics of information, *Journal of Political Economy*, 69 (3), 213-225

Stocking, G.W. and M.W. Watkins (1946) *Cartels in Action*: *Case Studies in Business Diplomacy*, New York: Twentieth Century Fund

Sutton, J. (2002) *Marshall's Tendencies, What can Economist's Know?*, Cambridge, Mass: MIT Press

Swedberg, R. (1991) *Joseph A. Schumpeter*: *The Economics and Sociology of Capitalism*, Princeton, NJ: Princeton University Press

Teece, D.J. (2012) *Strategy, Innovation and the Theory of the Firm*, Cheltenham: Edward Elgar

Teece, D.J., G. Pisano and A. Shuan (1997) Dynamic capabilities and strategic management, *Strategic Management Journal*, 18 (7), 509-533

Trenberth, L. (2009) *Principles of Organization and Management*, Los Angeles: Sage

Toraubully, W.A. (2018) Large market games, the law of one price and market structure, *Journal of Mathematical Economics*, 78, 13-26

Vives, X. (1999) *Oligopoly Pricing: Old Ideas and New Tools*, Cambridge, MA: MIT Press

Walker, D.A. (1973) Edgeworth's theory of recontract, *Economic Journal*, 83 (321), 138-149

Walras, L. (1954) *Elements of Pure Economics, or The Theory of Social Wealth* (ed. W. Jaffe), London: Allen & Unwin

Watts, D.J. (2003) *Six Degrees*: *The Science of a Connected Age*, London: William Heinemann

Weintraub, E.R. (2002) *How Economics became a Natural Science*, Durham: NC: Duke University Press

Williamson, O.E. (1975) *Markets and Hierarchies*, New York: Free Press

Williamson, O.E. (1985) *The Economic Institutions of Capitalism*, New York: Free Press

Whitin, T.M. (1953) Classical theory, Graham's theory and linear programming in international trade, *Quarterly Journal of Economics*, 67 (4), 520-524

Witt, U. (1998) Imagination and leadership: The neglected dimension of an evolutionary theory of the firm, *Journal of Economic Behaviour and Organization*, 35, 162-177
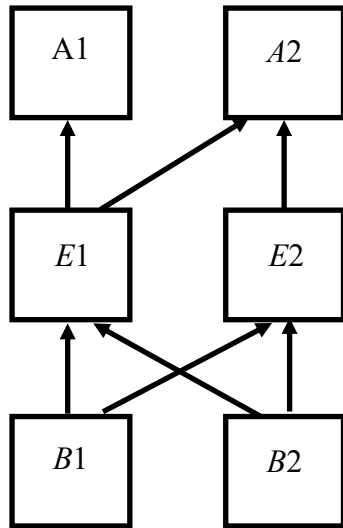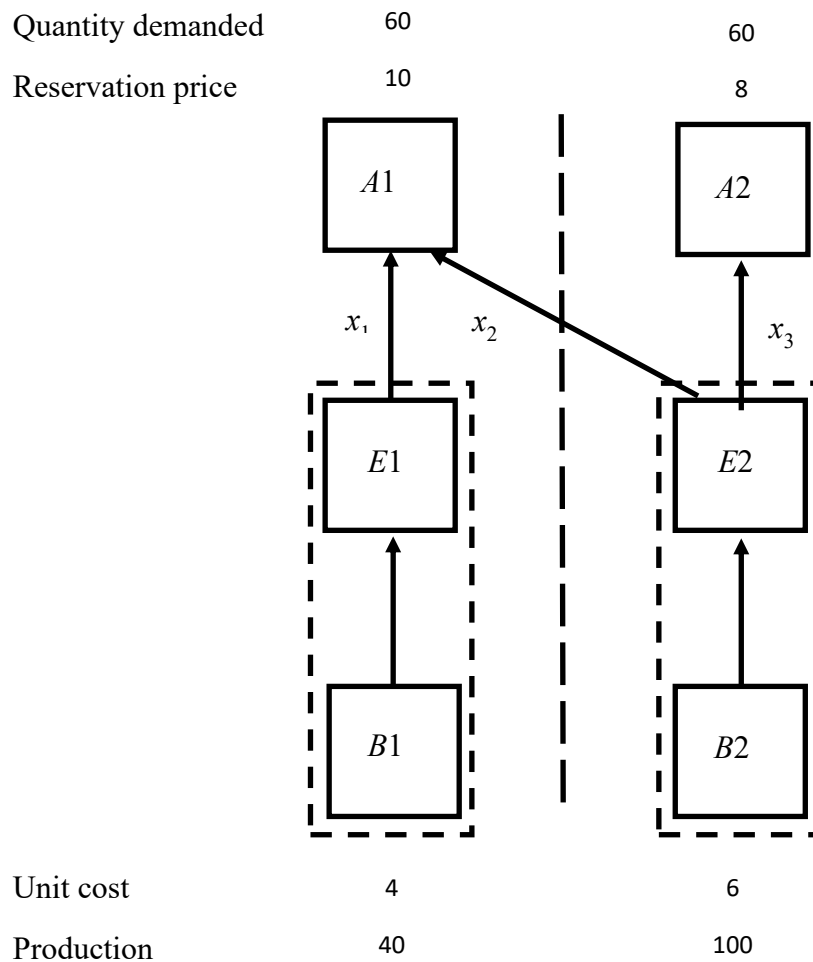
Figure 1

Typical configuration of a market

Figure 2

A simple segmented market with rivalry between entrepreneurs

Quantity demanded	60		60
Reservation price	10		8

$A1$ $A2$

x_1 x_2 x_3

$E1$ $E2$

$B1$ $B2$

Unit cost	4		6
Production	40		100

Note: There are no access costs.

Figure 3

Linear programming solution,

Illustrating the convex hull of feasible arrangements, with the equilibrium identified

Note: The number of feasible linkages determines the dimension of the convex hull. In the present case there are three available linages and so the diagram is three-dimensional. The equilibrium is identified by a black dot.
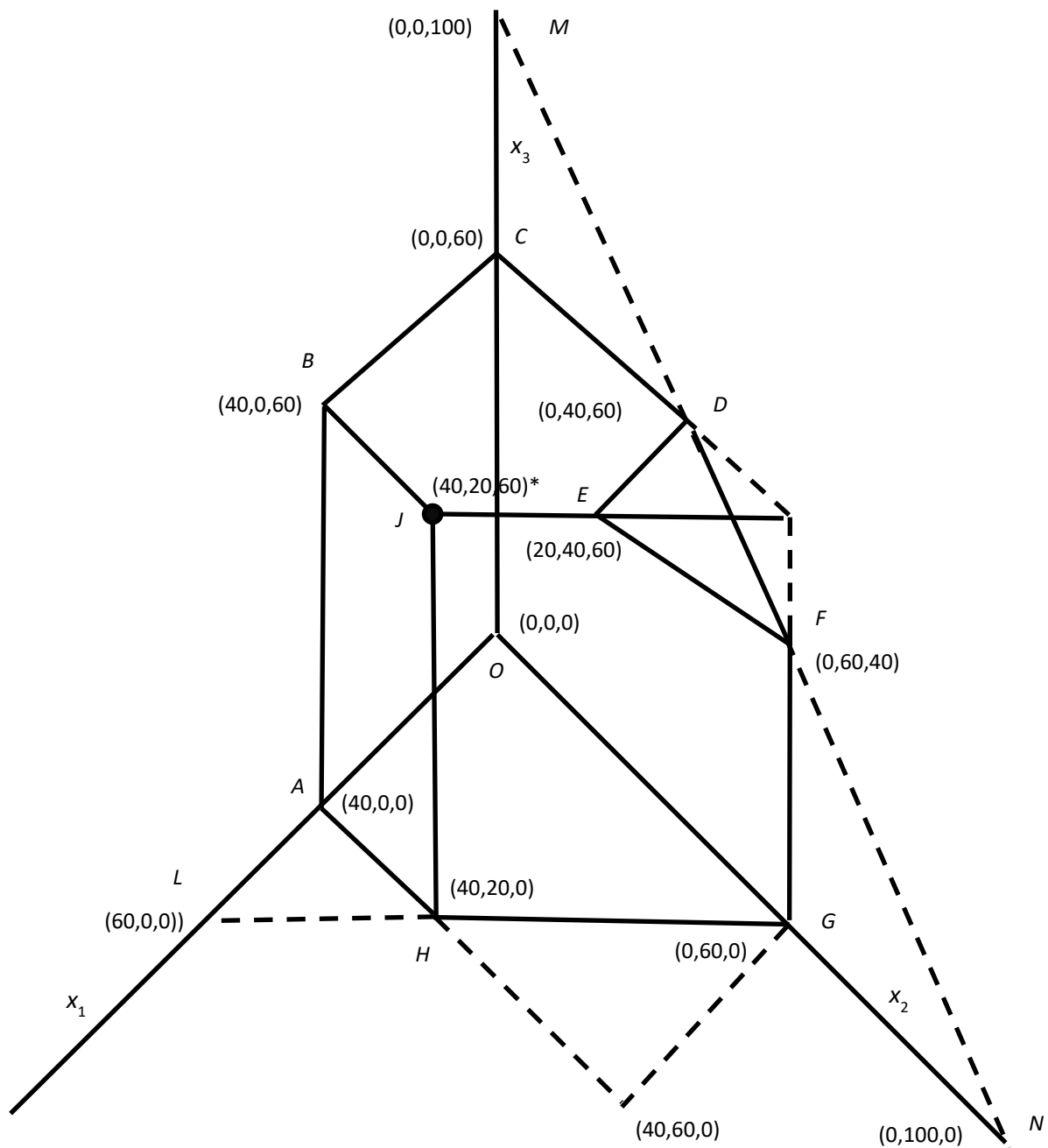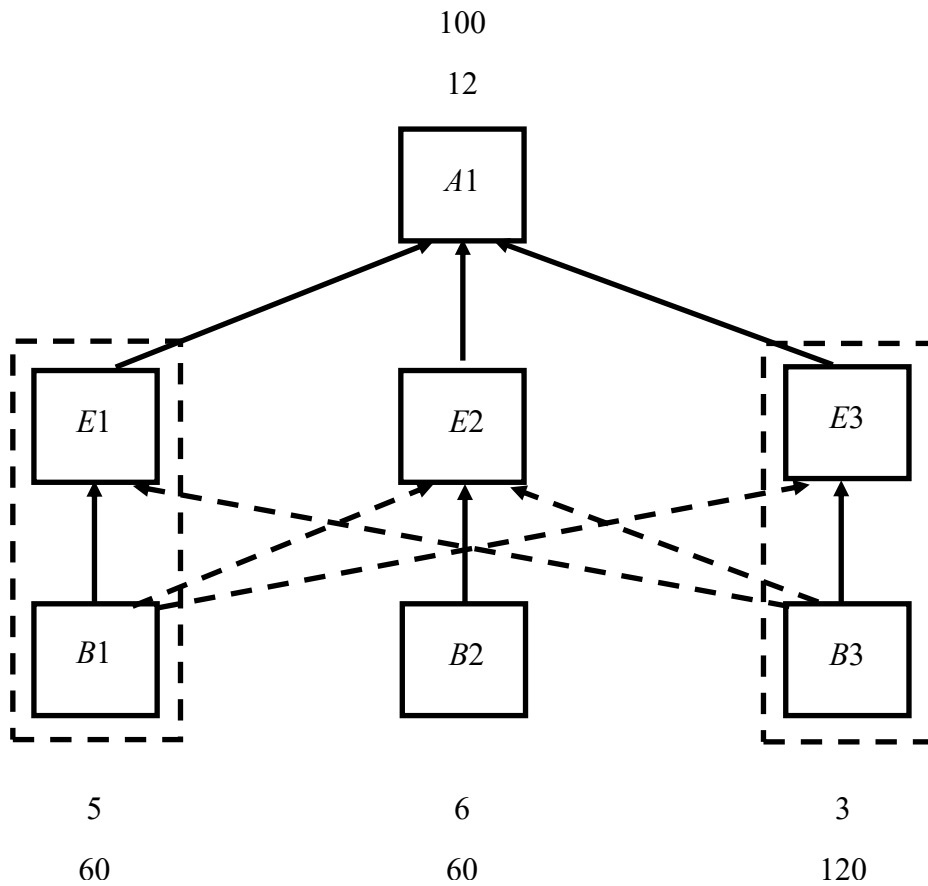
The diagram solves for quantities but not for prices.

Figure 4

Market evolution: Example 1

100

12

| A1 |

E1   E2   E3

B1   B2   B3

5   6   3

60   60   120

Key: A dashed rectangle identifies an integrated ownership unit

A dashed line indicates a potential linkage blocked by either *E*1 or *E*3

A solid line indicates a feasible flow of product
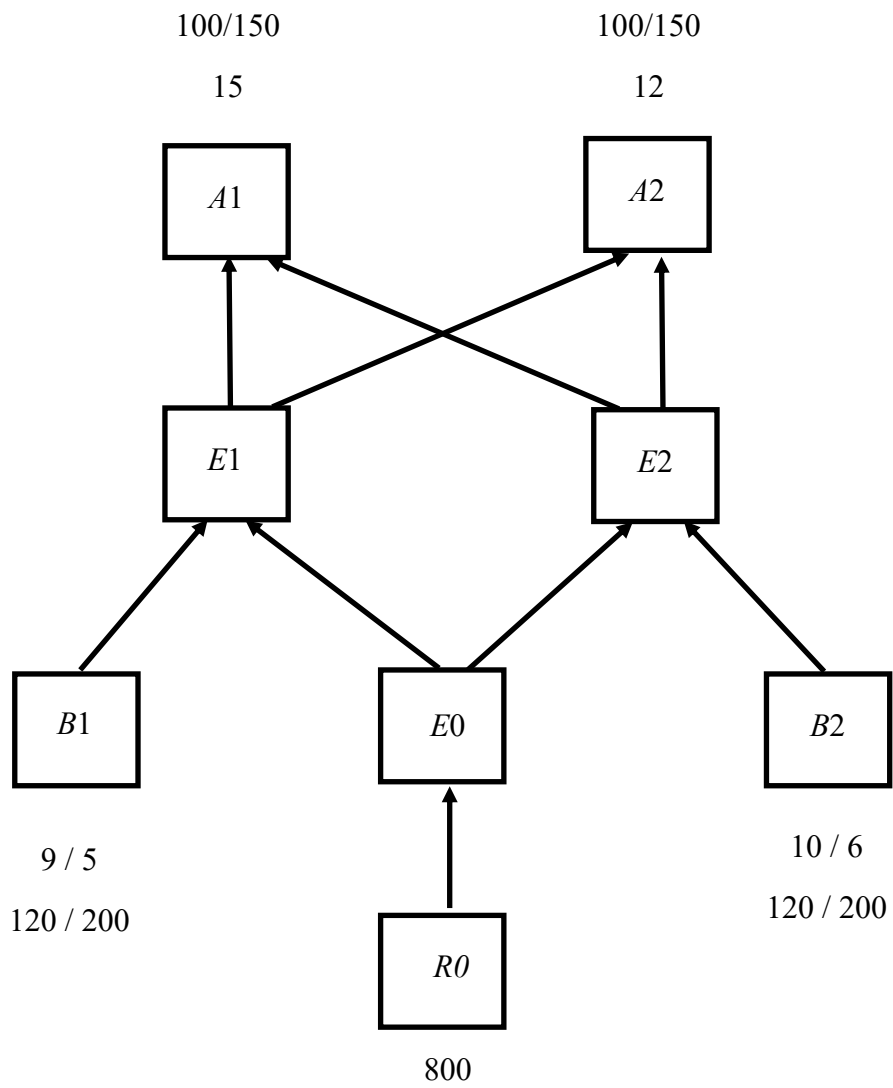
Figure 5

Market evolution: Example 2
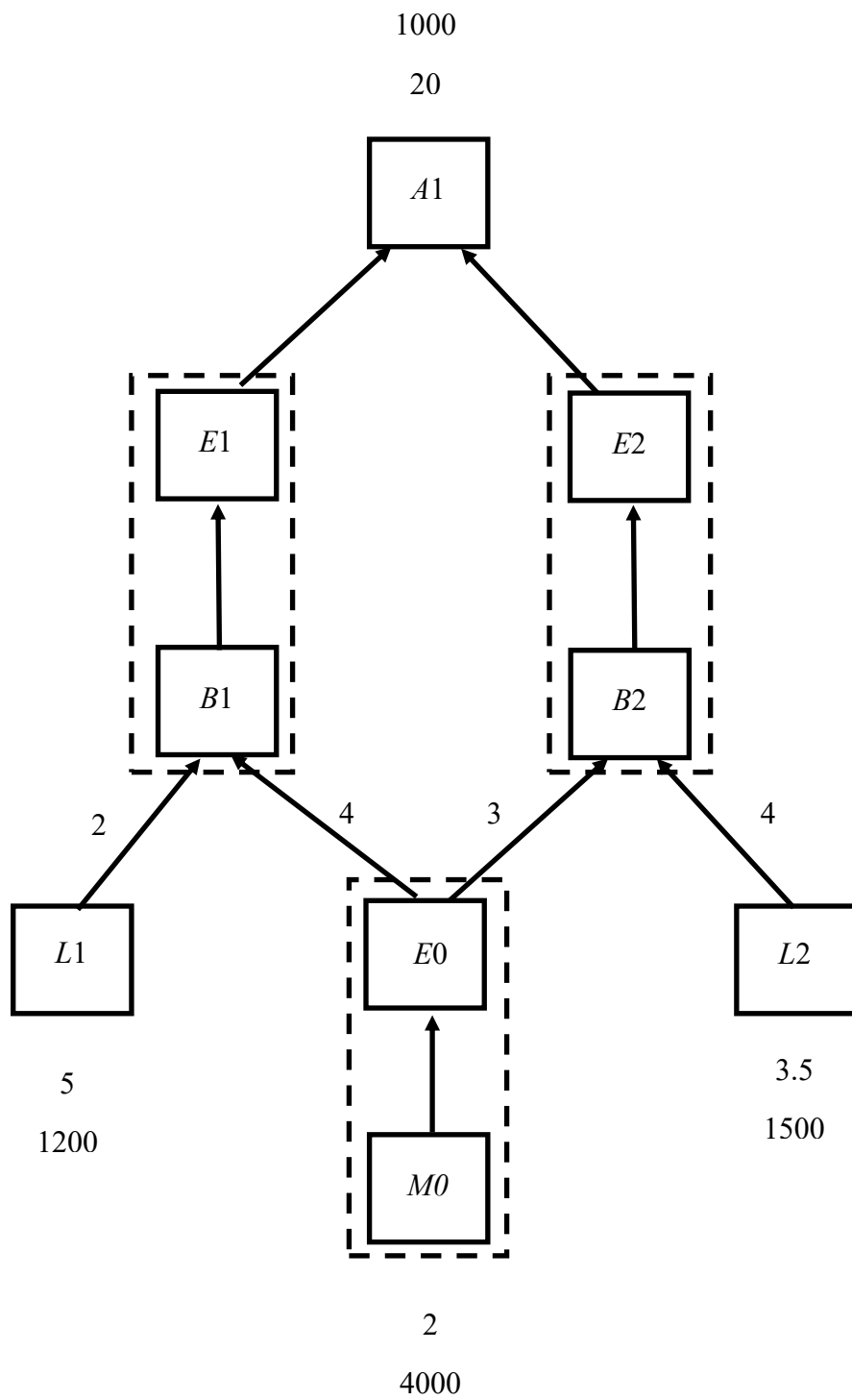
Figure 6

Market evolution: Example 3

Figure 7

Role of clubs in coordination