Marshall, Alfred (1842–1924)

John K. Whitaker
Edited by Steven N. Durlauf and Lawrence E. Blume

Abstract

English economist Alfred Marshall, founder of the Cambridge School of economics, was a leading and internationally prominent figure in the development of economic thought between 1870 and 1920. He played a significant role in professionalizing British economics, always stressing the social importance of wider economic understanding. His influential ideas on economic theory were conveyed primarily in his Principles of Economics (1890). Accounts of his life, career and general views on economics are followed by a more technical treatment of his contributions to various aspects of economic analysis. Guides to his writings and to the secondary literature on him are appended.

Keywords

Bernoulli hypothesis; Berry, A.; bimetallism; Böhm-Bawerk, E.; Bowley, A.; British classical economics; British Economic Association; Cambridge School; ceteris paribus; Chapman, S.; Clapham, J.; Clark, J.; combinations; comparative static method; conspicuous consumption; consumer surplus; cost and supply curves; cost of production; Cournot, A.; Cunynghame, H.; cyclical unemployment; deductive method; demand for money; demand theory; derived demand; differential rent; distribution theories (classical); Dupuit, J.; Edgeworth, F.; envelope theorem; equilibrium analysis; excess burden of taxation; external economies; Fay, C.; Flux, A.; Foxwell, H.; free trade; Gonner, E.; homo economicus; imperfect competition; institutions; internal economies; international trade (theory); interpersonal utility comparisons; Jenkin, H.; Jevons, W.; joint demand and supply; Keynes, J. M.; Keynes, J. N.; Lavington, F.; long-period equilibrium; long-period supply; Macgregor, D.; marginal cost pricing; marginal productivity theory; marginal utility of money; marginal utility theory; market demand elasticity; market interdependence; market value; Marshall, A.; Marshall, M. P.; Menger, C.; Methodenstreit; methodology of economics; Mill, J. S.; monetary theory; monopolistic competition; monopoly; multiple equilibria; multiplier; net product; Nicholson, J.; normal profit; normal value; offer curves; optimal taxation; Pantaleoni, M.; Pareto, V.; partial-equilibrium method; perfect competition; period analysis; Pigou, A.; poverty; Price, L.; principle of substitution; producer surplus; product differentiation; quasi-rent; Ramsey price; rent; representative firm; Ricardo, D.; Royal Economic Society; Sanger, C.; Say’s Law; self-interest; short-period equilibrium; Sidgwick, H.; Smith, A.; subjective theories of value; supply and demand; tariffs; temporary equilibrium; trade unions; utilitarianism; von Thünen, J.; Walker, F.; Walras, L.; welfare economics

JEL classifications

B31
Article

Alfred Marshall, Professor of Political Economy at the University of Cambridge from 1885 to 1908 and founder of the Cambridge School of Economics, was born in Bermondsey, a London suburb, on 26 July 1842. He died at Balliol Croft, his Cambridge home of many years, on 13 July 1924 at the age of 81. His magnum opus, *Principles of Economics* (1890a) evolved through eight editions in his lifetime, the final edition (1920) being most commonly cited today. It was one of the most influential treatises of its era and was for many years the Bible of British economics, introducing many still familiar concepts. The Cambridge School rose to great eminence in the 1920s and 1930s. A.C. Pigou and J.M. Keynes, the most important figures in this development, were among Marshall’s pupils.

Marshall’s biography and career are outlined initially, after which descriptions are given of his views on the social setting, aims and methods of economics, and his intellectual debts to others. An analysis of his fundamental ideas on theories of value and distribution, which were mainly set out in *Principles*, follows, after which his contributions to monetary and international-trade theory are considered briefly. A final section provides additional documentation and general suggestions for further reading. Also, some of the more technical sections have attached to them brief ‘bibliographic notes’ offering suggestions for further exploration. All bibliographic references lacking an author’s name are to works by Marshall, and the bibliographic details of all his cited publications can be found in the list of ‘Selected works’ below. The bibliographic details for all cited works written or edited by others are listed in the concluding ‘Bibliography’.

Biography and career

Marshall grew up in the London suburb of Clapham, being educated at the Merchant Taylors’ School where he showed academic promise and a particular aptitude for mathematics. Eschewing the more obvious path of a closed scholarship to Oxford and a classical education, he entered St John’s College, Cambridge, in 1862 on an open exhibition. There he read for the Mathematical Tripos, Cambridge University’s most prestigious degree competition, emerging in 1865 in the exalted position of Second Wrangler, bettered only by the future Lord Rayleigh. This success ensured Marshall’s election to a Fellowship at St John’s. Supplementing his stipend by some mathematical coaching, and abandoning – doubtless because of a loss of religious conviction – half-formed earlier intentions of a clerical career, he became engrossed in the study of the philosophical foundations and moral bases for human behaviour and social organization. In 1868 he became a College Lecturer in Moral Sciences at St John’s, coming to specialize in teaching political economy. By about 1870 he seems to have committed his career to developing this subject, seemingly ripe for reform, and helping to transform it into a new science of economics.

For several years he laboured persistently to develop and refine his economic ideas, and to deepen his understanding and grasp of both the existing economic literature and the economic reality that was its subject matter. In 1875 he visited the United States to probe economic conditions, and throughout his life he was tireless in his efforts to master the practicalities of the economic world. Prior to 1879 his publications were meagre. He had embarked on a book on international trade and problems of protectionism in the mid-1870s, and before that he had worked out
many of his distinctive theoretical ideas in the form of short essays, many now reproduced in Whitaker (1975). But the only part of this material to be made public was four chapters from the theoretical appendices for the proposed international-trade volume. In 1879 Henry Sidgwick had these printed for private circulation under the title *The Pure Theory of Foreign Trade: The Pure Theory of Domestic Values* (1879a). (An amplified version together with surviving portions of the text of the abandoned trade volume is also reproduced in Whitaker, 1975.) The year 1879 also saw the publication of Marshall’s first book, *The Economics of Industry* (1879b), written jointly with his wife Mary Paley Marshall.

Mary Paley had been one of the first group of students at Newnham Hall (later Newnham College) where Marshall, an early supporter of the informal scheme of Cambridge lectures for women, taught her political economy. Their marriage in 1877 required Marshall to give up his Cambridge position under the celibacy rules then in force. He found a new livelihood as principal of the recently established University College, Bristol, where he also became Professor of Political Economy. There *The Economics of Industry* was brought to completion and published by the house of Macmillan, which continued as Marshall’s publisher thereafter. Ostensibly an elementary primer, this book contained the first general statement of Marshall’s emerging theories, and a considerable sophistication lay beneath its deceptively simple surface. Together with the powerful *Pure Theory* chapters published by Sidgwick, a few copies of which circulated outside Cambridge, *The Economics of Industry* marked Marshall as a rising star in the economics firmament. With the death of W. S. Jevons in 1881, he moved into the public eye as the leader in Britain of the new scientific school of economics.

The duties of the Bristol principalship proved irksome to Marshall, especially as the college was struggling financially. He was anxious to proceed with his writing, having by 1877 conceived the plan for the book that was to become the *Principles*. His frustrations were increased by the onset in 1879 of a debilitating illness, diagnosed as kidney stones, which restricted his activities. He was persuaded to continue as principal until 1881, when he resigned both posts at the college. The next year was spent travelling, with an extended sojourn in Palermo, and it was in this year that composition of the new book began in earnest.

At Bristol, Marshall had got to know well Benjamin Jowett, the famed Master of Balliol, who was one of the governors of the struggling college. It was probably by Jowett’s generosity that Marshall was able to return to Bristol in 1882 as Professor of Political Economy. And it was doubtless at Jowett’s instigation that the Marshalls moved to Oxford in 1883, when a Balliol lectureship became vacant on the unexpected death of Arnold Toynbee. Marshall had considerable success as a teacher in Oxford and appeared settled in for an indefinite stay. But an ‘Oxford School of Economics’ was not to be. The sudden death of Henry Fawcett, who had been Professor of Political Economy at Cambridge since 1863, opened up the irresistible prospect of a return to Cambridge and a position with great potential for academic leadership. Marshall, the dominant candidate, was duly elected in December 1884, holding the chair until 1908, when he resigned to devote himself entirely to writing.

In many ways Cambridge’s inviting prospects were to prove illusory. Economics was taught as part of the Historical and Moral Sciences Triposes, but neither avenue provided a supply of able interested students, nor was there much scope for advanced work. Marshall struggled for many years, with limited success, to increase the scope for economic teaching. But it was not until 1903, with the
establishment of a new Tripos in Economics and Politics, that his goal was achieved. Even then, few resources were made available by the university and colleges for the teaching of economics, and the staffing of the new Tripos relied heavily on Marshall’s willingness to support two young lecturers from his own pocket. The flowering of the new school came about mainly after his retirement, but the seeds were certainly planted by his efforts.

Absorbed in the struggle for his own subject, Marshall took relatively little part in general university affairs. Indeed, his rather obsessive personality and proneness to magnify details would have made him ineffectual as a university statesman even if he had aspired in that direction. But he did play a prominent part in the successful campaign of 1896–7 against the granting of Cambridge degrees to students of the women’s colleges – this despite his wife being at the time a lecturer at Newnham. He was not opposed to women’s education, indeed had been a warm supporter in his early days, but was vehemently opposed to the assimilation of women into an educational system designed for men.

But the dominant fact in Marshall’s life after his return to Cambridge, and certainly the aspect of greatest interest to posterity, is his long struggle to give adequate written expression to the stores of economic knowledge and understanding he had accumulated. The demands of teaching and administration left him little time or energy for sustained composition during term time and it was in the jealously guarded long vacations, usually spent away from Cambridge on the south coast of England or in the Tyrol of Austria, that the only real progress could be made. By 1887 the book commenced in 1881 had grown into a projected two-volume treatise. He hoped to complete the first volume in time for it to appear in the autumn of that year with the second volume appearing by 1889. In fact, the first volume (1890a) appeared as the Principles of Economics, Volume One, only in July 1890, when it was received with great and immediate acclaim and established Marshall firmly as one of the world’s leading economists. The second volume never appeared. It was to have covered foreign trade, money, trade fluctuations, taxation, collectivism and aims for the future – a tall order!

Marshall struggled for the next 13 years with his intractable second volume, meanwhile spending much time on substantial, but not very substantive, recastings of the first volume in new editions of 1891, 1895 and 1898, and in preparing a digest of it to replace the earlier Economics of Industry which he had come to dislike intensely. (The digest, 1892, appeared under the title Elements of the Economics of Industry, Volume One. Like the earlier work it included material on trades unions that was never incorporated into Principles.) By 1903 much material had been accumulated for the second volume, but the scope was becoming unmanageable as Marshall became increasingly preoccupied with problems of trusts, trades unions, international trade, and comparative economic development, and increasingly concerned with matters of pure theory. In that year, partly from the impetus of writing a private memorandum on trade policy for the use of the then Chancellor of the Exchequer, and partly because the tariff controversy was at full heat, Marshall was tempted into writing a short topical book on foreign trade questions, intending to publish it speedily. But this project too grew unmanageably in his hands. In 1907, the preface to the fifth edition of Principles (the last major rewriting) announced the abandonment of the proposed continuation and promised instead a volume, already partly in print, on ‘National Industry and Trade’, to be followed soon by a companion volume on ‘Money, Credit and Employment’ (Guillebaud, 1961, vol. 2, p. 46). To reflect this change, the title of the sixth and subsequent editions of
Principles was changed to Principles of Economics: An Introductory Volume.

Retirement in 1908, at the age of 66, freed Marshall to concentrate on these projects, but progress continued to be slow. He appears to have suffered from recurrent dyspepsia and high blood pressure, necessitating a strict regimen and limiting his ability to work. But the more fundamental problem was that the world kept changing and the increasingly realistic and factual tone of his enquiry called for incessant recasting and revision. Nothing had been completed by the time war broke out in 1914, and then much rewriting was required to take into account the radical changes that were transforming the world economy and its post-war prospects. At last, when Marshall was 77 years old, Industry and Trade (1919), his second masterpiece, finally appeared. It was a magisterial, largely factual, consideration of trends in the British and international economy and of future economic prospects. But, lacking an obvious theoretical skeleton, it has not received from economists the kind of attention lavished on Principles, although interest in it is now beginning to stir among historians of economics.

In its final form, Industry and Trade was narrower in scope than had been intended earlier, while the proposed book on ‘Money, Credit and Employment’ still remained to be written. Over the next four years, by a remarkable effort, and despite rapidly waning powers, some of the mass of accumulated raw material remaining was pulled together in Money, Credit and Commerce (1923). This contains Marshall’s fullest treatment of the theories of money and international trade, but it is an imperfect pastiche of earlier material, some dating back almost 50 years.

In the last months of his life, Marshall toyed with the occasional writings and the memoranda and evidence for governmental enquiries that he had prepared at various stages during his career, with the hope of editing them for publication in book form. This was not to be, but his plan was largely fulfilled after his death in two books sponsored by the Royal Economic Society (Pigou, 1925; Keynes, 1926).

Judged by what might have been, Marshall’s authorial performance after 1890 was a sorry one, marked by repeated procrastination and inconstancy and by chronically over-optimistic expectations. The mantle of leadership that he had assumed on Jevons’s death had proved a heavy one. Both temperamentally and by virtue of his acknowledged position as the doyen of British economists, Marshall was compelled to attempt the magisterial and to denigrate the kind of forceful direct essay of which he was eminently capable.

As Cambridge professor and unquestioned leader of British orthodox economists, Marshall could hardly avoid becoming a public figure whose pronouncements carried more than a personal weight. His consciousness of this, and of the precarious public standing of economics, as well as his own temperament, made him peculiarly reluctant to enter into public controversy, although he would on occasion fire off a letter to The Times on some issue of the day. He served as an expert witness for several government enquiries and was an influential member of the Royal Commission on Labour of 1890–94. As President of Section F of the British Association in 1890 he took the formal lead in the movement to found the British (later Royal) Economic Association, but he was not a prime mover. Indeed, he was not a clubbable or organizational man and relied on others to further whatever goals he desired for economics and the economics profession at large. But neither was he a recluse. Balliol Croft received a continuing stream of visitors, ranging from working class leaders to distinguished foreign economists, while students or young colleagues were always welcomed and offered generous advice mixed with exhortation.
Although able students interested in economics were in short supply, Marshall did over the years teach and influence several students who were to make contributions to the subject. From the early Cambridge period H.S. Foxwell, H.H. Cunynghame, J.N. Keynes and J.S. Nicholson might be mentioned. The Oxford period brought L.L.F.R. Price and E.C.K. Gonner, while the period as Professor in Cambridge produced, among others, A. Berry, A.W. Flux, C.P. Sanger, A.L. Bowley, S.J. Chapman, A.C. Pigou, J.H. Clapham, D.H. Macgregor, C.R. Fay, and, last but not least, J.M. Keynes.

The undoubted fact of Marshall’s professional leadership of British economics calls for some explanation. He was far from suited to such a role by temperament, and his fussiness and inflexibility could be irritating. For example, Sidgwick, J.N. Keynes, and Foxwell, the most important of his early allies in Cambridge, were all eventually alienated. Marshall’s success can be attributed partly to sheer persistence. As in the case of the new Tripos, he had a clear idea of what he wanted to accomplish and worried away at it until he exhausted the opposition and was allowed to have his way. But it must also have been due to the lack of any alternative. The relevant question is not ‘Why Marshall?’ but ‘Who else?’ Economics was rapidly evolving as a profession around the turn of the 20th twentieth century, creating a leadership vacuum. Leadership was unlikely to emanate from outside Oxford, Cambridge or London, but F.Y. Edgeworth at Oxford was perhaps the last man capable of meeting the need, while E. Cannan at the new London School of Economics, although more suited than Marshall to the hurly-burly of professional politics, was too much the perennial critic and iconoclast to fill the bill. Moreover, whatever Marshall’s foibles, the sheer power of his intellectual vision, his international standing as Britain’s leading economic thinker, and his ability to inspire an impressive flow of budding scholars, all conspired to make him the only feasible contender.

**Marshall’s views on the social setting, aims and methods of economics**

Marshall saw economics as concerned with those aspects of human behaviour open to pecuniary influences and sufficiently regular and ubiquitous to permit statements of broad scope and some persistence. While maintaining, especially in earlier work, that some heeded moral imperatives might be impervious to pecuniary considerations, he conceded that most behaviour lay within the ambit of the measuring rod of money. On the other hand, he emphasized that motivation was not merely a matter of pursuing pecuniary self-interest, even if broadly conceived to include interests of family and friends. He was anxious to lay the ghost of *homo economicus* and emphasized the human desires to obtain social approbation or distinction and to enjoy the pleasures of skilful activity. He saw actors as diverse as captains of industry and sculptors driven more by the joys of creative activity and the striving for the regard of peers than by the desire for material acquisition.

As well as not being pecuniary maximizers in any narrow sense, individuals were for the most part seen as imperfect optimizers. The working classes, especially, often lacked the knowledge and foresight to judge their long-term interests. Marshall’s actors were not imbued with complete knowledge of their environment but had to acquire knowledge slowly, and often painfully, through experience. Nor were they endowed with fixed desires and an intrinsic, unchanging character. Indeed, character and preferences evolved as individuals were exposed to new possibilities
and chose to enter into new activities. The workplace, in particular, was an important moulder of character. Self-improvement and character development induced by environmental changes, planned or unplanned, both figured largely in Marshall’s world view. He believed that social institutions, such as land tenancy practices, were pliable and ultimately moulded themselves into conformity with the individual interests involved, rather than presenting a permanent constraint on mutually desired accommodations. (For this he was taken to task by his most vehement critic, W. Cunningham, who denied the applicability of modern economic theory to medieval practices – see Cunningham, 1892.) But institutional change must be slow, slower even than changes in individual character and wants, because informal customs and tacit agreements are hard to change. Thus, while the institutions and informal understandings and prohibitions that constrain and mould economic behaviour might ultimately be endogenous they will often be ill adapted to current circumstances and thereby act as an independent constraint on the pursuit of mutually desired accommodations. Institutions, in the broad sense, are important and not always socially rational constraints on individual action.

Marshall was impelled to economics because ‘the study of the causes of poverty is the study of the causes of the degradation of a large part of mankind’ (1920, p. 3). For the bulk of the population, mired in poor living and working conditions, little progress in habits, aspirations and self-esteem could be expected without prior improvement in economic conditions. Such improvement was socially important not so much for its own sake, at least once the pangs of immediate want were assuaged, but because of its instrumental role in permitting and stimulating improvement in the quality and character of the population. What Marshall really valued was not improvement in the standard of living but the enhancement of the standard of life that this improvement made possible. And he entertained little doubt about what constituted a qualitative improvement here, even though – or perhaps because – his values may seem quite parochial and culture-bound.

Economic improvement required appropriate institutions, incentives and attitudes, and would be threatened by wide-scale government intrusions into economic affairs, although some forced income redistribution could be tolerated. But even if economic conditions were improved, the full yield of social betterment would be garnered only if enlarged consumption were turned to ennobling and horizon-expanding channels (rather than, say, to strong drink), involved a due consumption of beneficial leisure, and was accompanied by healthier and less stultifying conditions of working and town life. The government had a guiding role to play here. But even more important would be the assistance and example of employers and the upper and middle classes, who must first rid themselves of a frequent propensity to showy and ostentatious consumption and excessive materialism. The working-class leaders and skilled artisans who had already raised their own standard of life had an important leadership role too. Voluntary individual efforts to assist the rise of the underprivileged must rest on an adequate understanding of economic consequences. For this, as well as to secure an informed electorate, the diffusion of sound economic knowledge was an essential and integral element in the process of socio-economic transformation. Economics thus was itself a noble activity of high importance for the future of mankind.

The broad view of the economy suggested by the foregoing is of a complex evolutionary process of combined economic, social and individual change in which each individual’s abilities, character, preferences and knowledge develop jointly, along with social institutions, markets and the technologies of production and
communication. The pursuit of self-interest, broadly conceived, is ubiquitous in directing this evolutionary process, but is subject to inertia, ignorance and limited foresight, not to mention individual mutability.

Unfortunately, Marshall was able to bring little formal analysis to bear on this general ‘biological’ vision of the economy and could only evoke it descriptively. It might be true that ‘the Mecca of the economist lies in economic biology rather than in economic dynamics’ (1920, p. xiv). Nevertheless, the only available analytical tools were those of classical mechanics, tools that Marshall’s early mathematical training had equipped him to employ skillfully. In fact, chief reliance had to be put on that branch of classical mechanics dealing with statics. Dynamics, beyond a few qualitative applications, required more precise information than was likely to be available. Perforce then, much of Marshall’s formal analysis, like that of W.S. Jevons or Léon Walras, was based on simple assumptions of individual optimization and market equilibrium, taking preferences, technology and market institutions for granted. Such provisional or tentative ‘statical’ treatments could often be valuable. Indeed Marshall viewed them as indispensable for the correct analysis of many questions. But he was always anxious to stress that the analysis was preliminary, and perhaps of only transitory validity. This awareness made him impatient of over-elaboration, so that, for example, he showed no interest at all in pushing the statical approach to its logical conclusion in the general equilibrium analysis of the stationary state. For him, equilibrium analysis was an indispensable but rough and ready instrument that needed to be employed with due caution and a continuing awareness of its limitations in the face of a complex ever-evolving reality. It was only a tool and did not itself constitute concrete knowledge.

Marshall had no great profundity as a philosopher of science and had little patience with metaphysics: ‘in a sense ... he held no views on method’ (Coase, 1975, p. 27). Marshall’s discussions of methodology largely reflect the philosophical presuppositions of his day. His method was in the general deductive tradition of John Stuart Mill, but he sought to emphasize the relativity of particular theories, as contrasted with the universality of the general theoretical ‘organon’ or economists’ toolbox. Anxious to present a public image of the unity of economics in the face of the Methodenstreit among economists in the late 19th century, he attempted to maintain an uneasy balance on method, decrying extended chains of deductive reasoning but denying the possibility of purely inductive inference unguided by a coherent conceptual framework. Economics had room for specialists in both deductive and inductive methods, but both must ultimately be co-workers. Assumptions must be selected with close regard to the facts of the case and potential disturbing causes must be kept prominently in mind and due allowance made for them. J.N. Keynes described Marshall’s analytical method as ‘deductive political economy guided by observation’ (1891, p. 217n) and Keynes’s chapter ‘On the Deductive Method in Political Economy’ (1891, pp. 204–35) is perhaps as good a rationalization of Marshall’s method as one can find.

**Intellectual debts**

The intellectual background to Marshall’s work in economics was established in the 1860s, partly in his stringent mathematical training, but perhaps more importantly in the heady mixture of utilitarianism, evolutionism and German idealism which he eagerly imbibed in the years immediately following his graduation. He seems to have
started on economics from J.S. Mill’s *Principles of Political Economy* (1848), moving on to the classic works of Smith and Ricardo. At a fairly early stage, probably around 1868, he discovered Cournot’s *Récherches* (1838), which provided examples of the application of mathematics to economic questions. Acquaintance with J.H. von Thünen’s work, which influenced Marshall’s distribution theory, must have come somewhat later, in the early to mid-1870s. During the 1870s and early 1880s Marshall also read widely on economic development and socialism, including much literature in German, the only foreign language he mastered thoroughly. After that, his reading seems to have been concentrated mainly on factual and practical matters. Once his own theoretical views had crystallized, he appears to have been reluctant to do more than attempt to explain and clarify them to others, and to have taken remarkably little interest in new theoretical issues or in the theoretical ideas of others.

In many ways, the list of Marshall’s denials of theoretical indebtedness is more remarkable than that of his acknowledgments. He claimed to have developed his ideas on consumer surplus before learning of anticipations by J. Dupuit and H. Fleeming Jenkin. The grudging attitude to W.S. Jevons’s marginal utility theory shown in his review (1872) of Jevons (1871), although subsequently relaxed, was never replaced by any acknowledgement of indebtedness. He showed little or no interest in the work of Walras, gave meagre credit to Carl Menger, whose work must have become known to him by the early 1880s, patronized Pantaleoni and Böhm-Bawerk, largely ignored Pareto, and so on. Even in the case of Edgeworth, one of his few intimates, Marshall felt that undoubted theoretical powers were guided by an unreliable judgement and refused to follow Edgeworth’s subtle elaborations far. In fact, the only major theorist of the day to command Marshall’s entire admiration and respect was J.B. Clark, and even here there was no acknowledgement of serious indebtedness. This tendency to denigrate the work of his contemporaries was matched by an equally strong tendency to overvalue the achievements of the British Classical School led by A. Smith, D. Ricardo and J.S. Mill. For one reason or another – perhaps a personality quirk, perhaps an effort to boost the public esteem of economics – Marshall was prone to exaggerate the intellectual continuity and maturity of his subject – see O’Brien (1990) on this.

A growing interest in wider intellectual influences on Marshall in his formative years 1865–70 has been sparked by the publication and analysis of his early philosophical manuscripts (Raffaelli, 1994, 2003), especially a paper entitled ‘Ye Machine’ that outlines a mechanism capable of learning new routines from experience, thus freeing its limited learning ability to gradually establish new and higher level routines, and so on. It appears that Marshall’s ambitions in these early years lay in the area of ‘psychology’ or perhaps better in the ‘philosophy of mind’. Whether the world lost more than economics gained from his switch to economics remains an open and perhaps insoluble question. But it does appear that the pattern of a sequential routinizing of new methods, continually leading to new levels of individual or organizational complexity, continued to play a significant part in Marshall’s economic thought. More generally it is clear that he read philosophical literature widely in his formative years: Kant, Hegel, H. Spencer, and others. But whether and how these sources influenced his economic thought remains uncertain, partly because evidence is slight or absent.
Demand theory

So far the discussion has remained on a very general level, dealing with broad aspects of Marshall’s life and work. At this point there begins a much more detailed and technical consideration of various aspects of his theoretical contributions, commencing with his demand theory. Marshall’s treatment of the theory of demand is sketchy and incomplete, concentrating on the demand for a single commodity, or commodity group, against a loosely defined background. A utility-maximizing individual’s utility is defined by \( u(x) + w(y) \) where \( x \) is the individual’s consumption of the particular good \( X \), while \( y \) is the individual’s expenditure on all other goods. This expenditure is measured in money of constant purchasing power: that is, deflated by a general price index. How this index is defined and whether, as seems appropriate, the price of \( X \) is excluded from it, is left unclear. Such money can be treated as a composite good, \( Y \), and \( y \) can be regarded as the amount of this composite good consumed. If \( m \) is the individual’s initial endowment of \( Y \), then \( y = m \) whenever \( x = 0 \), while if \( X \) can be freely purchased at a fixed price of \( p \) units of \( Y \) per unit of \( X \) then \( x \) and \( y \) must satisfy the constraint \( px + y = m \). Marshall assumes that the utility functions \( u(x) \) and \( w(y) \) have positive but diminishing marginal utility so that \( u'(x) > 0 > u''(x) \) and \( w'(y) > 0 > w''(y) \), where single and double primes are used to denote first and second derivatives. The maximum expenditure, \( e \), that the individual is willing to make to secure \( x \) units of \( X \) is implicitly defined as a function \( e(x, m) \) by \( u(x) + w(m - e) - w(m) = 0 \). Providing that \( x \) and \( y \) are both positive, the rate at which \( e \) increases with \( x \) is \( u'(x)/w'(m - e) \) by the implicit function theorem. This ratio would be the demand price for the \( x \)th unit of \( X \) if all previous units had been acquired at their corresponding demand prices: that is, if the individual had faced perfect price discrimination in exchanging \( Y \) for \( X \). Alternatively, if the individual had been able to obtain any amount of \( X \) at a fixed per unit price, \( p \), the resulting demand function \( x(p, m) \) for \( X \) would be implicitly defined (given \( x \) and \( y \) are both positive) by the first-order condition \( u'(x) - pw'(m - px) = 0 \). Partial differentiation of \( x(p, m) \) shows that \( x \) falls as \( p \) increases, while an increase in \( m \) increases both \( x \) and \( y \); thus, the Giffen possibility of an increase in \( p \) increasing the quantity of \( X \) demanded is excluded. But an increase in \( p \) may lower or raise the value of \( px \), so that demand for \( X \) may be price elastic or price inelastic at a given \( p \).

The possibility of buying at a fixed price rather than facing perfect price discrimination creates a consumer surplus of \( e(x(p, m), m) - px(p, m) \). This is the additional amount that could have been extracted by perfect price discrimination for all units up to the price-taking optimal one. That this surplus is positive follows from the fact that every infra-marginal unit of \( X \) acquired creates a surplus utility when the individual faces a fixed price (since \( u'(x) > pw'(m - px) \) for each such \( x \) but no surplus when the individual is faced with perfect price discrimination.

Marshall’s mathematical notes (1920, pp. 838–42) on his general case are obscure and puzzling. Doubtless he felt this case was too dependent on unobservables to be of much practical value. He therefore emphasized the special case in which the marginal utility of money is treated as a constant. The rationale offered is that an individual’s ‘expenditure on any one thing . . . is only a small part of his whole expenditure’ (1920, p. 842). This simplifies \( e = e(x, m) \) above to \( e = u(x)/w'(m) \) while \( x(p, m) \) is now defined implicitly by \( u'(x)/w'(m) = p \). At the \( x \) value defined by the latter equation, consumer surplus arising from the ability to buy
any amount of X at the per-unit price $p$ can be expressed in utility terms as $u(x) - xu'(x)$ or in money terms as $u(x)/w'(m) - xu'(x)/w'(m)$. These formulae are exactly analogous to the standard formula for Ricardian land rent, with the first term the output obtained on a piece of land from the application of $x$ doses of variable input, each dose remunerated at the common marginal product. Partly because of this analogy, Marshall used the term ‘consumer rent’ rather than ‘consumer surplus’ prior to 1898.

Although priority must go to Dupuit, Marshall’s simple concept of consumer surplus based on the assumption of a constant marginal utility of money has been influential. But he was well aware of the complications arising from variation in the marginal utility of money: ‘Strictly speaking we ought to take account of the fact that if he spent less on tea the marginal utility of money to him would be less than it is, and he would get an element of consumers’ surplus from buying other things at prices which now yield him no such rent’ (1920, p.842). Although such influences may be ‘of the second order of smallness’ they raise the more disquieting issue of assessing the overall welfare effects of changes that affect many markets simultaneously. On this Marshall had little to say: ‘the task of adding together the total utilities of all commodities, so as to obtain the aggregate of the total utility of all wealth, is beyond the range of any but the most elaborate mathematical formulae’ (1920, p.131n.). It was a task he chose not to pursue. Apart from generalizing for the possibility that a certain quantity of good X might be indispensable, Marshall elected not to develop his demand theory further, or even to generalize it to incorporate utility functions that were not additively separable (1920, p. 845). It is clear that each commodity in turn might take the spotlighted role of good X and that in certain circumstances simultaneous consumer surpluses for several goods might be added (1920, p. 842). An unpublished early manuscript note from the 1870s on the theory of taxation (Whitaker, 1975, vol. 2, pp. 285–305) had advanced matters considerably further by working formally with the maximization of utility under a budget constraint, but this lead was not followed up in print and some of its lessons for welfare economics were apparently forgotten. Principles gave a clear intuitive account of the consumer’s overall optimization problem (1920, pp. 117–23), but failed to connect it to the resulting interrelated set of demand functions for the various goods consumed. Indeed, it is clear that for positive purposes Marshall was willing to treat market demand functions in a quite pragmatic way, admitting, for example, close substitutes or complements and the Giffen exception, all inconsistent with the simple formal theory set out above. In judging this, it must be borne in mind that consistency and generality of ‘stational’ analysis were not Marshall’s real goal. Rather, ‘fragmentary stational hypotheses are used as temporary auxiliaries to dynamical – or rather biological – conceptions’ (1920, p. xv).

The market demand for a good that is offered to all actual or potential buyers at the same given price is of course obtained as a function of that price by summing the amounts demanded at that price by all the consumers. A sufficient but not necessary condition for market demand to fall as price increases is that each individual’s demand decreases. The now familiar concept of market demand elasticity – proportional quantity change divided by proportional price change – was first introduced by Marshall, although several authors had come close to the idea previously. It appeared without flourish in (1885c), and appeared more prominently in Principles. But Marshall himself made relatively little use of it.

Bibliographic note: Marshall’s treatment of demand is essentially contained in (1920, pp. 92–137, 838–43). An influential, although controversial, interpretation of

Production and long-period competitive supply

In deriving the long-period supply curve of a commodity in *Principles*, Marshall envisages production as organized by firms, typically family businesses. Each firm strives to minimize its production costs, substituting one productive factor or production method for another according to the ‘Principle of Substitution’. In its simpler forms this involves marginalist adjustment to bring relative marginal value products into line with relative marginal costs. But more generally, the Principle of Substitution is akin to a natural selection process, being ‘a special and limited application of the law of survival of the fittest’ (1920, p. 597). Marshall’s firms do not have costless access to a common production function, but must grope and experiment their way to cost-reducing modifications. The long-period supply curve is defined for a given state of general scientific and technical knowledge. But each firm must explore this to some extent anew.

Although the distinction is not entirely clear – distinctions seldom are for Marshall – two polar cases may be distinguished within his theory of long-period competitive supply. These will be referred to as the ‘agricultural’ and the ‘industrial’ cases. The former is much the more straightforward and involves an industry in which production is relatively simple, internal economies of scale are minimal, and the product is homogeneous and easily marketed. The optimal firm size is small, and management is sufficiently routine to need no exceptional ability to keep a firm operating efficiently. As the overall market expands, new firms may be added, but changing composition of the population of firms is not an essential feature of this case.

The long-period supply price per unit of output at which such an industry can supply any quantity of output must just cover the cost of maintaining that level of output indefinitely. That is, it must just suffice to pay all the inputs (including management) needed to produce that level of output in a cost-minimizing way at rates that just ensure that the requisite input quantities will continue to be forthcoming indefinitely. In the case of skilled workers, in particular, the rate must just suffice to induce parents to apprentice new workers to the industry at a rate exactly offsetting the attrition through retirement and other causes. Similarly, the return to fixed capital must just suffice to induce replacement of the existing stock of fixed assets, while the return to management must keep up the necessary replacement flow of managers. On the other hand, the return to land services must just suffice to prevent these services from migrating elsewhere, replacement not being necessary. As the level of industry output being considered is increased, the supply price will probably rise, mainly because of the need to pay a higher return to land so as to attract a greater supply from other uses, but perhaps also because of the need to pay more for rare natural talents that, like land, must be attracted in greater quantity from other uses, not being capable of replication through education and training. Such a tendency for long-period supply price to rise with output may be mitigated though seldom eliminated by substitution against inputs whose supply price is rising, and by possible external economies that increase each firm’s efficiency by influences that
depend, not on its own output, but on the entire industry’s output. A tendency for supply price to rise with output will imply that infra-marginal units of those inputs whose supply prices are rising receive rents, since all units will be remunerated at the rate necessary to induce continuing supply of the marginal unit. In the absence of external economies (or diseconomies) the total rent or producer surplus generated will be the ‘triangular’ area above the supply curve. That is, it will be

\[ R = xg(x) - \int_0^x g(v) \, dv \]

where \( g(x) \) is the supply price of output quantity \( x \), an increasing function of \( x \). This result does not apply in the presence of external economies. In later editions of the Principles, Marshall introduced the device of the ‘particular expenses curve’ (1920, pp. 810–12) to display rent in such a case, but this ex post construction does not give an independent basis for determining rent.

It is clear that the long-period supply curve of an industry depends on the general economic background against which the industry is assumed to operate. As is the case with demand, Marshall does not consider this background in detail. He assumes prices to be expressed in money of constant purchasing power and recognizes on occasion that there may be close interrelations between two industries (for example, they may compete for the same specialized land). He also recognizes that ‘a theoretically perfect long period must give time enough to enable not only the factors of production of the commodity to be adjusted to the demand, but also for the factors of production of those factors of production to be adjusted and so on’ (1920, p. 379n.), and that this leads ultimately to the assumption of a stationary state. But he is not willing to follow this route far and is content in general to take the supply conditions of the factors of production for granted when analyzing long-period price determination.

In the ‘manufacturing’ case, to which we now turn, the product is differentiated, marketing is difficult, and each firm must build up and retain goodwill and a customer connection for its own specialized product. There are substantial internal economies of scale in production and successful management calls for business ability of a high and rare character. In this environment, a family business may be built up by an exceptional founder, but this build-up must be slow because of the difficulty of establishing a market and perhaps also because of constraints on financing. And when the founder passes on, his successors are unlikely to have equal talents or even the lesser talents required to prevent the firm’s business from languishing. By the third generation of succession, the firm is likely to expire. Even a joint stock company (a case added rather as an afterthought) is likely to ossify into bureaucratic stagnation, and presumably the same is true of family businesses that rely on paid managers. Thus, the typical firm in the manufacturing case passes through a finite life cycle, and the industry is comprised of a population of such firms at various phases of the life cycle, some in the early expanding phase, others in decline.

The long-period supply price at which such an industry can supply a specified level of output must now be regarded as an index of the prices of all the different firms’ products. It must meet all the conditions required in the agricultural case. Thus, the price must allow for a continuing replacement flow of the various types of workers (including managers) and fixed assets, as well as the retention of the necessary ‘land’ services. But now there must also be a surplus sufficient to induce a
replacement flow of new firms – a supply of ‘business organization’ that will just suffice to replace the expiring firms and keep the age distribution of firms constant.

Industry equilibrium does not require each firm to be in an unchanging equilibrium any more than the trees in the proverbial forest. A new firm will be established if the prospective earnings over the expected life cycle appear to justify the cost and trouble involved. The firm’s initial earnings are likely to be negative as it slowly builds up its technical expertise and market connections, but these early losses can be regarded as investments to be recouped in the later stages of the firm’s prospective life cycle.

It is here that Marshall’s ‘representative firm’ enters the picture. It is best regarded as a parable that avoids the need to consider the entire distribution of firms. By definition, the long-period supply price of any level of industry output is the average cost of the representative firm at that level of output. Industry-level magnitudes may then be regarded as if they were generated by a fixed number of unchanging representative firms rather than by the actual heterogeneous body of ever-changing firms. In other words, the manufacturing case may be treated as if it were an agricultural case populated by representative firms only. Such arguments add nothing conceptually and are prone to confuse, although it might be noted that Marshall believed an acute well-informed observer could select an actual firm that was close to being representative in this sense.

The average cost and size of the representative firm will change as industry output changes. There are two main reasons for this. A larger industry output is likely to generate more external economies, lowering the costs of every firm. But more importantly, the larger industry demand is, the easier it will be for a new firm to build up a market, and so the larger the size to which firms will grow before they begin to decline. This will bring about greater access on average to unexhausted internal economies of scale, again leading to lower costs on average. For both these reasons, long-period supply price is likely to decline as a larger industry output is considered, even though the opportunity cost of obtaining greater supplies of land services and rare natural talents may rise. Again, the particular expenses curve may be used to display the producer surpluses or rents accruing to such scarce factors at any given level of industry output, but the relationship of this family of curves to the long-period supply curve is tenuous and complex. Rent obviously cannot be represented by a ‘welfare triangle’ above the supply curve when the latter is falling.

The conception of competition in Marshall’s manufacturing case is much closer to later ideas of imperfect or monopolistic competition than to modern notions of perfect competition. Products are differentiated and firms are not price takers, but face at any time downward-sloping demand curves in their special markets. Even if the difficulties of rapidly building up a firm’s internal organization can be overcome, the resulting enlarged output cannot be sold at a price covering cost – even granted substantial scale economies in production – without going through the slow process of building up a clientele and shifting the firm’s particular demand curve. The time this takes is assumed to be considerable relative to the duration of the firm’s initial vitality. But in some cases the difficulties of rapid expansion may be overcome. They may not have been very severe, as when different firms’ products are highly substitutable, or the firm’s founder may have unusual genius. In such cases the industry will pass into a monopoly or be dominated by a few, strategically interacting firms, or ‘conditional monopolies’ as Marshall termed them.

Marshall’s reconciliation of persisting competition with increasing returns and falling supply price is complex and problematic, but it does not depend in any
essential way on scale economies being external to the firm. The concept of external economies is one of his significant contributions, although his treatment of it can hardly be called pellucid. But it was added more for verisimilitude than because it was theoretically essential to the structure of his theory.

The issues surrounding Marshall’s representative firm, and the problem of reconciling the persistence of competition with the presence of unexhausted internal economies of scale, continue to receive attention among historians of economic thought but no definitive reading has yet been attained, or perhaps ever will be. The account given above is well supported by Marshall’s text, but as is often the case with Marshall, elements of ambiguity and vagueness remain.


Price determination and period analysis

The long-period supply curve for any good indicates for each market quantity the least price at which that quantity will continue to be supplied indefinitely. The long-period equilibrium price and quantity are determined by the intersection of this supply curve with the market demand curve, assumed to be negatively sloped, that indicates the highest uniform price at which any total quantity can be sold. In the agricultural case, equilibrium will be unique as the supply curve slopes positively. But in the manufacturing case, the supply curve, as well as the demand curve, may have negative slope, so that multiple equilibria can occur. Equilibrium is adjudged locally stable if demand price is above (below) supply price at a quantity just below (above) the equilibrium quantity. The intuitive justification for this is that the actual price of any available quantity is determined by the demand price, while quantity produced tends to increase (through both expansion of existing firms and entry of new firms) whenever an excess of market price over supply price promises high profits, while it tends to decrease in the opposite case.

This stability argument is sketchy and, in any case, there still remains the question of exactly how a new long-period equilibrium is attained following some change, such as a permanent shift in the demand curve. One possibility would be to consider explicitly the adjustment process through time, but Marshall preferred to approach the problem by another route – his period analysis, one of his most memorable and lasting contributions. (His passing claim (1920, p. 808) that the long-period supply curve may not be reversible, supply price depending upon past-peak output as well as current output, is something of an exception to this generalization. It appears to rest on some restriction of the degree of downward supply adjustment, and so not to involve a true long-period analysis, or else to invoke a kind of learning by doing that once attained is not readily lost.)

Period analysis is Marshall’s most explicit and self-conscious application of the comparative-static, partial-equilibrium method with which his name will always be associated. As he observed,
the most important among the many uses of this method is to classify forces with reference to the time which they require for their work; and to impound in *Caeteris Paribus* those forces which are of minor importance relatively to the particular time we have in view. (Guillebaud, 1961, vol. 2, p. 67)

Which forces or variables are to be hypothetically frozen or impounded, and which are to be determined by the requirements of equilibrium (an equilibrium contingent upon the contents of the *ceteris paribus* pound, of course), should be determined pragmatically in each case with the aim of focusing on the features deemed dominant in that case. As a general rule, those forces should be impounded which move very slowly, or else bounce around very rapidly, relative to the length of ‘the particular time we have in view’. This is well illustrated by Marshall’s example of a fish market, where the focus may be on the determinants of price over a few days, a few months, or several years, or even decades (1920, pp. 369–71). As an expositional matter, however, and also to embody distinctions of wide (but not universal) applicability, Marshall emphasized three broad cases. Temporary or market equilibrium analysis proceeded on the assumption of a fixed stock of output already available or in the pipeline. Short-period normal equilibrium analysis permitted output to be varied, but not the stock of productive ‘appliances’ available to produce that output. ‘Appliances’ must be taken here to cover skilled labour and business organization as well as fixed capital assets, so that the existing set of firms is to be taken as given. Finally, long-period normal equilibrium, which has already been considered, allows the stock of appliances, as well as the level of output, to be freely varied. In this case equilibrium incorporates the conditions necessary for inducing an exact replacement flow of each kind of appliance, including a replacement flow of new firms in the manufacturing case.

Temporary equilibrium for a perishable commodity is simply a matter of selling off the existing stock. Marshall recognizes the possibility of ‘false trading’ – sales at a non-equilibrium price – but argues that (a) this will not affect the eventual price if the marginal utility of money is constant, and (b) price will quickly settle close to the uniform price that would just clear the market if used in all transactions. With a storable good there is the further speculative possibility of holding back supply for future sale, and this gives expected future cost of production an indirect role in influencing current market price. Cost of production already incurred is an irrelevant bygone, however.

In short-period normal equilibrium, output is adapted to demand within the constraints set by the fixed supply of available ‘appliances’. High demand will raise equilibrium output, but only within the limits possible by working existing appliances more intensively or pulling in versatile unspecialized labour and land from elsewhere. Low demand will lead to low utilization of appliances, perhaps idleness of some, and migration of unspecialized inputs to elsewhere. In the agricultural case a firm will change output until marginal prime or variable cost equals market price. In the manufacturing case, a fear of spoiling the future market or invoking retaliation from competitors tends to make a firm’s output more responsive to variation in market price, and hence to make market price less responsive to demand shifts. Otherwise, the two cases are similar, both involving a fixed population of firms and a rising supply curve.

The return received by an appliance will often exceed the minimum necessary to induce its operation at the chosen intensity (its prime cost) and this excess is a ‘quasi rent’. To the extent that land and rare natural talents are immobile in the short
period, or less mobile in the short period than the long, their returns too will often have a quasi-rent element. Otherwise, they will receive only differential rents, though often at rates differing from their long-period values. It should be stressed that the concepts of quasi-rent and differential rent are relative to a specific use. The prime cost necessary to retain an input in this use may itself include rent or quasi-rent when viewed in the context of a more inclusive set of alternative uses. Thus, from the viewpoint of all possible uses in the economy, the return to any factor in fixed supply is entirely a rent or quasi-rent (the latter if fixity is only short-period).

Marshall paid little attention to the possibility that forces similar to those constraining the adjustment of supply when time is limited might also operate on the side of demand. Thus the same considerations underlie the market demand curve whether it is coupled with a temporary, short-period or long-period supply curve. In each case, market equilibrium price and quantity are determined by the intersection of the appropriate demand and supply curve. The stability of temporary equilibrium is directly asserted. The stability of short-period equilibrium depends on the same quantity-adjustment argument invoked for long-period equilibrium, but since the short-period supply curve is always positively sloped, uniqueness and stability are assured.

The theory of short-period normal equilibrium was designed as a tool for analysing unemployment and economic fluctuations in the never-completed second volume of *Principles*. But it also has use in explaining adjustment to a permanent disturbance. Suppose, for instance, that an industry is in long-period equilibrium when a permanent shift in demand occurs. The immediate or short-period effects can be analysed by freezing output and stocks of appliances at their initial levels. Insight into the actual adjustment through time can then be obtained by appropriately changing the output level assumed in the temporary equilibrium, so that movement of temporary equilibrium towards short-period equilibrium can be traced out as output, but not stocks of appliances, adjusts. Similarly, the levels assumed for the stocks of appliances in this short-period equilibrium can be allowed to change and the movement of short-period equilibrium towards long-period equilibrium traced out. Such arguments are now a staple of elementary pedagogy. They clearly require additional assumptions about the adjustment of output and the way in which investment or disinvestment in appliances proceeds, and are only a poor and ambiguous substitute for an explicit dynamic analysis. But such ‘statical’ procedures, although imperfect, may, in Marshall’s words, be ‘the first step towards a provisional and partial solution in problems so complex that a complete dynamical solution is beyond our attainment’ (Pigou, 1925, p. 312).

Marshall’s period analysis, and more generally his partial-equilibrium approach to price determination, was designed in large part as a usable tool for the analysis of concrete issues. Its longevity amply testifies to its usefulness in this respect. But it was also meant to serve the more doctrinal purpose of clarifying the respective roles utility and cost of production play in determining value. The aim was to show that the greater the scope for supply adjustment permitted in the definition of equilibrium, the more dominant the supply side influence on price becomes. This doctrinal goal helps to account for the rather heavy weight given to long-period analysis in *Principles*. For, as Marshall recognized, its value as a tool of applied analysis is seriously qualified by the fact that ‘violence is required for keeping broad forces in the pound of Caeteris Paribus during, say, a whole generation, on the ground that they have only an indirect bearing on the question in hand’ (1920, p. 379n). That is, there is no good ground for assuming that background forces such as technology and...
tastes will remain constant for the length of time required for long-period equilibrium to be practically relevant. For concrete analysis of problems of such long duration it will often be necessary to transcend the period analysis, with its reliance on statical equilibrium, and undertake directly an analysis of secular change, of which Book 6, Ch. 12 of *Principles* on the ‘General Influence of Economic Progress’ (1920, pp. 668–88) offers the main example, but not a very impressive one.

In emphasizing the role that cost of production plays in the determination of long-period value, Marshall was not content to rest on money costs of production but sought to go behind these costs to the real costs – the efforts and abstinences – for which in a non-coercive economy the money costs are recompense. In doing so he purported to follow Ricardian tradition, but is more plausibly viewed as attempting to place the newer subjective value theories in broader (but still subjective) focus. Just as the price paid by a consumer serves as a measure of marginal utility, with a consumer surplus gained on infra-marginal units, so the unit price received by a worker or saver measures the real cost or disutility at the margin, with a producer surplus on the inframarginal units of effort or abstinence. But, as Marshall recognized, the parallel holds imperfectly in the long period when workers must be regarded as produced means of production as well as final consumers and cost bearers. In particular, parental sacrifice for raising and training offspring obtains little or no direct pecuniary reward.


**Normal value and normal profit**

Implicit in the preceding discussion are Marshall’s conceptions of normal value and normal profit. Normal value is defined as the value that would result ‘if the economic conditions under view had time to work out undisturbed their full effect’ (1920, p. vii). It is contrasted with market value, which is ‘the actual value at any time’ (1920, p. 349). Normal value is hypothetical, resting on a *ceteris paribus* condition, its role being to indicate underlying tendencies. The normal value of a commodity may approximate its average value over periods sufficiently long for the ‘fitful and irregular causes’ (1920, pp. 349–50) that dominate market value to cancel out, but this should not be presupposed automatically outside a hypothetical stationary state.

The distinction between normal and market value is closely related to the distinction between natural and market value found in the work of Smith and the classical economists. In 1879 Marshall had identified normal value with ‘the results which competition would bring about in the long run’ (1879b, p. vii), but in *Principles* he switched to the view that ‘Normal does not mean Competitive’ (1920, p. 347) and admitted any kind of regular influence so long as it was sufficiently persistent. The economic forces hypothetically permitted to achieve full mutual accommodation could now be chosen appropriately for each case. In particular, the distinction between short-period and long-period normal (or ‘sub-normal’ and ‘true-normal’ in earlier editions) was emphasized.
Profit was viewed by Marshall as the residual income accruing to a firm’s owner, a return on the investment of the owner’s own capital and recompense for the pains of exercising ‘business power’ in planning, supervision and control. Normal profit is essentially an opportunity cost, the minimum return necessary to secure the owner’s inputs to their current use, or rather to accomplish this for an owner of normal ability. Marshall presumes that there is a large and elastic supply of versatile actual or potential owner managers of normal ability. In long-period equilibrium each of these must just receive the same normal rates of return on investment and exercise of business power whatever the line of business. However, those who are exceptional may do better, essentially by exerting greater business power.

These common rates of normal return are simultaneously determined, along with the normal returns to other kinds of effort and abstinence, by Marshall’s macroeconomic theory of the long-period determination of factor incomes (see below). Although it is the case that profits are a residual, rather than a contractually agreed amount like other incomes, this difference is immaterial in long-period equilibrium. In particular, a long-period equilibrium analogy between ordinary wages and the normal earnings of business power is stressed. Normal profit is a necessary element in the costs that underlie the long-period normal supply curve, but actual profit is a quasi-rent or producer surplus for shorter periods.

Normal profits are a return to ‘business power in command of capital’ and compensate for three distinct elements: ‘the supply of capital, the supply of the business power to manage it, and the supply of the organization by which the two are brought together and made effective for production’ (1920, p. 596). The combined compensation of the latter two components comprises ‘gross earnings of management’, the return to the second component being ‘net earnings of management’. In long-period equilibrium, the normal return to the first element is imputed at the market interest rate on default-free loans, and that to the second component at the rate paid to hired managers performing comparable tasks. The residual third element, the return to ‘organization’, is most straightforwardly interpreted as an extra return on owned capital equivalent to the premium for default risk, or ‘personal risk’, that would have to be paid on borrowed capital. In the manufacturing case, the annual level of normal profit for each firm in an industry must be interpreted as the annualized equivalent of the expected stream of returns that is just sufficient to induce an individual of normal ability to found a firm in the industry rather than divert energies and capital elsewhere. Normal ability here is defined relative to other potential founders of firms, a group already exceptional relative to the population as a whole. By construction, such normal profits must be earned by the representative firm.


**Welfare economics**

To serve as a tool of welfare economics, monetary measures of consumer surplus, producer surplus and rent must be aggregated over individuals. But how are the resulting sums to be interpreted? Marshall’s very limited and proximate attempts at
formal welfare arguments are carried out within a utilitarian framework, for which the goal is maximizing aggregate utility. He implies that interpersonal utility comparisons are possible in principle and that utility functions will be similar for all members of any group that is homogeneous in terms of mental, physical and social attributes. Within such a group, the marginal utility of money will be the same for two individuals having the same income, and lower for the richer of two individuals with differing incomes, on the assumption in each case that both individuals face the same trading opportunities. A postulated government action may impose gains and losses on various individuals that can be measured and aggregated in money-equivalent terms. But how can these measures be translated into statements about aggregate gains and losses of utility? Marshall emphasizes two special cases. First, if the gains and losses are both proportionately distributed over income classes in exactly the same way, then net aggregate gain (positive or negative) in money will serve as an ordinal index for the net aggregate gain in utility. A corollary of this is that if two alternative actions affecting the same group have the same relative distributions of gains and losses over income classes then the alternative yielding the greater net aggregate gain in money must have the greater net aggregate gain in utility. Second, if some change makes for a zero net aggregate change in money terms, but the gains accrue to individuals of lower income than those bearing the costs, then the aggregate net utility gain must be positive – a warrant for certain redistributive policies. In other cases he sees that careful assessments of the marginal utility of money to the various injured and benefited groups would be needed, assessments that could be used to transform monetary gains and losses into utility measures. He toys (1920, pp. 135, 842–3) with using the Bernoulli hypothesis on the relation between wealth and utility as a basis for such calculations, but gives little indication as to how assessments might be made in practice.

Marshall’s best known and most successful foray into formal welfare analysis was his proof that total welfare might be increased by using the proceeds of a tax on an ‘agricultural’ industry to subsidize a ‘manufacturing’ industry. All comparisons involved long-period equilibria and relied on the validity of aggregated money-equivalent measures of gains and losses. He demonstrated that the gain in consumer surplus in the expanded decreasing-cost manufacturing industry might exceed the combined loss in consumer and producer surplus in the contracted increasing-cost agricultural industry. No formal account was taken of a possible gain in producer surplus in the manufacturing industry as this merely makes the argument hold a fortiori. The crucial point in this argument, as Marshall recognized, is that producers are not harmed by ‘a fall in price which results from improvements in industrial organization’ (1920, p. 472). It is immaterial whether the improved organization of the enlarged manufacturing industry is due to external economies or to internal economies resulting from an increase in the size of the representative firm. Contrary to much subsequent opinion, Marshall’s tax-subsidy argument is not necessarily dependent upon external economies.

Another significant, but overlooked, welfare analysis provided by Marshall was that of a monopolistic public enterprise in a situation where taxation involves an excess burden (1920, pp. 487–93, 857–8). In the absence of this excess burden Marshall proposes that the enterprise seek to maximize ‘total benefit’, the sum of net profit and consumer surplus. This implies marginal cost pricing, since the area below the demand curve and above the marginal cost curve is maximized when the two curves intersect. But, given that taxation involves an excess burden, it may be desirable to augment tax receipts from monopoly revenue if the sacrifice of
consumer surplus is small. Marshall proposes the alternative goal of maximizing ‘compromise benefit’, the sum of consumer surplus and monopoly revenue when the latter is in effect multiplied by the marginal cost of raising a unit of government revenue from other sources. Maximization of compromise benefit leads to the setting of what has come to be termed a ‘Ramsey price’.

The two examples of welfare analysis just described proceed within a partial equilibrium framework, treating each industry as negligible compared to the entire economy and regarding the marginal utility of money as approximately constant to each individual. Marshall’s rather fragmentary remarks on optimal tax systems, income redistribution and the ‘doctrine of maximum satisfaction’ cannot be restricted in this way, and so raise serious unresolved analytical difficulties. On the other hand, his tax-subsidy argument was a valid counterexample to arguments that competition must lead to a social optimum, or that optimal indirect tax systems must involve uniform tax rates. It must also be borne in mind that utilitarian welfare economics was for Marshall only a first step towards a more evolutionary analysis of possible modes of improving the physical quality and the values and activities of mankind.


Interrelated markets and distribution theory

Marshall was anxious to emphasize the interdependence of markets and introduced his treatment of joint and composite demand and supply largely for this purpose. A group of goods is jointly supplied if all are outputs of a single productive activity and jointly demanded if all are inputs. On the other hand, a particular good is compositely supplied or demanded if it is provided or acquired by several distinct productive activities. Marshall’s formal treatment of joint demand and supply proceeded on the general assumption that the products involved were consumed or produced in fixed proportions, as did his related analysis of the ‘derived demand’ for any one of several jointly demanded inputs – ‘derived’ since the demand for such inputs is derived from the demand for their joint product. The derived demand curve for a specific input can be constructed conceptually by supposing that its supply is perfectly elastic at an arbitrary price and that the markets for the output and all the inputs (including the specific input) adjust to equate quantity demanded to quantity supplied in each market. This gives a price quantity combination on the derived demand curve for the specific input. Other such combinations can be obtained by varying the arbitrarily chosen price and repeating the exercise, and so on. Marshall laid down four rules for inelasticity of derived demand. These were that the input should have no good substitutes, that the product it helps make should be inelastically supplied, that the input should account for only a small part of production costs, and that cooperating inputs should be inelastically supplied. Fixity of input ratios guaranteed the first condition, but the more general case was asserted rather than proven. The advantage of working with the derived demand curve for an
input is that it permits a more transparent analysis of the effects of changes in the supply conditions of the singled-out input.

The prime example of joint demand is the demand for productive inputs, and Marshall’s analysis of market interdependence was carried through more fully in this specific connection, the role of substitution among inputs receiving full acknowledgement. The principle of substitution ensured that input usage tended to be adjusted by firms so as to minimize the total production cost of any level of output. Thus, the value of the marginal product of an input (or the ‘net product’ as Marshall termed it) tended under competition to equal the unit price of the input. There has been some confusion about the relation between ‘net product’ and marginal product because the former allows usage of other inputs to adjust consequentially when the chosen input is increased while the latter does not. But, provided that the initial situation is cost-minimizing, the adjustment of other inputs (if small) has no effect on the change in output – an application of the envelope theorem. Marshall recognizes this explicitly (1920, p. 409n) and there is no good reason for refusing to classify him as a marginal productivity theorist.

Interdependence among input markets was further highlighted in the analysis of the competition of several industries for an input that is in temporarily or permanently fixed overall supply. A peculiarity of this last analysis was the insistence on excluding from the marginal cost of any industry the cost of bidding such fixed resources away from other uses. This is a perfectly legitimate application of the general envelope theorem: provided resource use is optimally adjusted, the marginal cost of increasing output will be the same whatever input or sub-group of inputs is increased. But Marshall’s insistence on asymmetry where there is really symmetry can be accounted for only by his desire to legitimate, and extend to quasi-rent, the classical doctrine that rent is price determined rather than a price-determining element of cost.

Marshall’s vision of market interdependence culminates in his treatment of income distribution, where he seeks to bring out the extents to which the interests of different factors of production are harmonious or conflicting. Distribution is determined by the interaction of the demands and supplies for the various inputs, the demands being essentially joint demands. Marginal productivity is a theory of input demand, not a complete theory of distribution, because the supplies of the various inputs cannot be viewed as fixed, at least in the long period. Indeed, in the long period the dominant influences on the prices of factors other than land are exerted by their supply conditions. The costs that then have to be met must ensure that various kinds of labour and capital continue to be replaced in their existing uses and quantities.

From an overall view ‘The net aggregate of all the commodities produced is itself the true source from which flow the demand prices for all these commodities, and therefore for the agents of production used in making them’ (1920, p. 536). This aggregate, ‘the national dividend’, is distributed among the factors of production. It is at once the aggregate net product of, and the sole source of payment for, all the agents of production within the country: it is divided up into earnings of labour; interest of capital; and lastly the producer’s surplus, or rent, of land and of other differential advantages for production. It constitutes the whole of them, and the whole of it is distributed among them; and the larger it is, the larger, other things being equal, will be the share of each of them. (1920, p. 536)
The share going to any class of inputs will depend upon the need people have for its services: ‘not the total need, but the marginal need’ (p. 536, italics original). But a complicating influence for distribution theory, although one ‘more full of hope for the future of the human race than any that is known to us’ lies in the fact that ‘highly paid labour is generally efficient and therefore not dear labour’ (p. 510). Influenced by F.A. Walker, Marshall was a strong proponent of the ‘economy of high wages’ argument that high wages increase labour efficiency, not perhaps immediately, but cumulatively over time and perhaps over generations: effects that transcend simple theorizing in terms of static equilibrium.

All the different productive factors cooperating in production have a common interest in increasing the size of the pie to be shared, the national dividend or income, but each factor has a selfish interest in restrictive practices that increase its own share, even if they reduce the size of the pie slightly. A prime question of social policy for Marshall is how these divergent incentives can be reconciled: how combined action by various groups, such as unions, can be prevented from assuming forms that, while perhaps individually beneficial to any one group in isolation, are certainly mutually harmful if undertaken by all.

Marshall here enters into macroeconomic forms of argument, and it is indeed true that he did toy with the formal specification of macroeconomic models of growth and distribution (see Whitaker, 1975, vol. 2, pp. 305–16). But, with this exception, it should be emphasized that his treatment of market interdependence fell far short of a full theory of general equilibrium on Walrasian lines. Even when formalizing market interdependence in the mathematical appendix to *Principles* (1920, pp. 846–56), he simply treated the demand or supply of each commodity as a function of nothing but the price of the commodity itself. The links between the generation of income in factor markets and the expenditure of that income in product markets were left quite vague. Again, it must be recalled that the development of comprehensive fully articulated equilibrium theories was not Marshall’s aim.


**Monopoly and combination**

Marshall’s analysis of price and output determination by a profit-maximizing monopolist, and of the effects of taxing such a monopolist, followed the lead of Cournot. The concept of marginal revenue was implicit in the mathematical statement, but Marshall’s chosen vehicle was geometrical. Curves of average revenue and cost, and of their difference, average net revenue, \( y \), (all functions of the quantity sold, \( x \)) were superimposed on a grid of iso-profit hyperbolae of form \( xy = \text{constant} \). Profit was maximized when the average net revenue curve touched the highest such iso-profit curve. Weighting consumer surplus into the maximand, as well as net revenue, gave rise to the welfare analysis of ‘compromise benefit’ already mentioned.

Monopoly analysis was applied to trades unions, with the use of the concept of the derived demand for an input. A union controlling a labour input for which
derived demand is inelastic can certainly raise wages – not only the wage rate but the total wages received – although at the price of unemployment of some members. Whether such a monopolistic restriction can be sustained for long is more doubtful, as there will be pressures both to enter the union and to evade its grasp by the relocation or reorganization of production.

A more problematic question was whether ‘labour’s disadvantage in bargaining’ meant that combined action by workers could raise wages, even without any restriction of labour supply. Marshall believed that it did, but emphasized that the result might be less capital accumulation by non-workers, an outcome that could harm workers eventually.

The extremes of monopoly and competition were both covered by the theory of normal value, even though the competition might be more akin to later concepts of imperfect or monopolistic competition than to any ideal form of perfect competition. But ‘normal action falls into the background, when Trusts are striving for the mastery of a large market’ (1920, p. xiv). The incidents, tactics and alliances of oligopolistic conflict defied reduction to a simple general theory. They were to have been considered in the uncompleted second volume of Principles and were to some extent covered by Industry and Trade. The latter’s treatment of entry-limiting behaviour by a ‘conditional monopolist’, who dominates the market but does not control entry, is of considerable interest in the light of much recent work on this class of problems.


Monetary theory

Marshall was in full command of previous British discussions of monetary issues, but not himself a major contributor to the development of monetary theory. His evidence before royal commissions in 1887 and 1899 showed an impressive mastery of monetary analysis, both domestic and international, and was minutely examined by successive generations of Cambridge students, serving for many years virtually as a textbook. But it was not until 1923, with the appearance of Money Credit and Commerce, that Marshall put forward his monetary views in a systematic way. By then these had not the novelty, nor he the vigour, to advance contemporary discussion.

Marshall’s most important contribution to monetary theory was to place the overall demand for money in the context of individual choices as to the fraction of one’s wealth to keep on hand as ready cash. This approach, set out clearly in a manuscript of the early 1870s (Whitaker, 1975, Vol. I, pp. 164–77), was developed by Marshall’s Cambridge successors, especially A.C. Pigou and F. Lavington, into what is termed the ‘Cambridge k’ approach. It laid the background for the treatment of the demand for money in J.M. Keynes (1936). On international monetary theory, Marshall espoused a form of purchasing power parity.
Marshall’s name is particularly associated with his proposals for ‘symmetalism’, the use of a fixed-weight combination of gold and silver as the monetary base, and for indexed contracts based on a ‘tabular standard of value’, or price index, to be maintained by the government. The former was offered as an improvement on fixed-ratio bimetallism, of which he was never more than a lukewarm adherent.

Marshall had interesting, if fragmentary, insights into business fluctuations and general unemployment, which he viewed as temporary disequilibrium consequences of credit market dislocations. These spilled over into general coordination failures, with unemployment in one market spreading to others by reducing demand in cumulative fashion – the germ at least of the multiplier concept. On the other hand, Say’s Law was maintained as an equilibrium truth of great importance. He saw the remedies for cyclical unemployment in the ‘continuous adjustment of means to ends, in such a way that credit can be based on the solid foundation of fairly accurate forecasts’, and in curbs on reckless inflations of credit that are ‘the chief cause of all economic malaise’ (1920, p. 710).


**International trade**

Marshall’s major contribution to international trade theory was his well-known geometrical analysis of the equilibrium and stability of two-country trade by means of intersecting offer curves. Each country’s offer curve indicated the number of bales of home goods it was prepared to exchange for a specified number of bales of foreign goods, demand being elastic or inelastic as an increase in the latter caused the former to increase or decrease. Possibilities of multiple and unstable offer-curve intersections were noted. The offer curves themselves were taken as data, although complex readjustments of production and consumption underlay them. The need for a separate theory of international trade was justified, in classical vein, by the supposed international immobility of factors of production that remain mobile domestically.

The main purpose of this theoretical apparatus was to examine the effects of tariffs. A country might gain by selfishly exploiting its monopoly power through restricting trade, and would certainly gain if trading equilibrium occurred on an inelastic portion of the foreign offer curve. But Marshall came to doubt increasingly the transferability of this result to a multi-country case, although admitting that it might apply to an export tax on an exceptional commodity (like British steam coal) lacking close substitutes and incapable of being produced elsewhere.

A related attempt to construct a theoretical measure of the ‘net benefit’ a country gains from foreign trade, analogous to the measures of consumer and producer surplus, was not entirely satisfactory as the partial equilibrium context had clearly been transcended.

On matters of concrete trade policy for Britain, Marshall was a firm but cautious adherent of free trade, even unilateral free trade, but became increasingly
concerned with the prospects for Britain’s position in the world economy. The discussion in Industry and Trade of the links between foreign competition and domestic industrial organization and structure reflected this concern.


A brief survey of Marshall’s writings with suggestions for further reading

The first editions of Marshall’s five books were (1879b), (1890a), (1892), (1919), (1923). Economics of Industry (1879b) had a new edition in 1881 and was reprinted with minor changes several times up to 1892. It is an important source for Marshall’s views on distribution theory, trades unions and business fluctuations. His magnum opus, Principles (1890a), had new editions in 1891, 1895, 1898, 1907, 1910, 1916 and 1920. The title was changed to its final form (as in (1920)) in the fifth edition. Principles is the basic source for Marshall’s views on the theories of value and distribution as well as his broader views on economics and social welfare. Since the rewritings between editions were substantial, the ninth variorum edition, edited by C.W. Guillebaud, Marshall’s nephew (Guillebaud, 1961), is essential for serious study. The first of its two volumes is a facsimile of the eighth edition of 1920. The second volume contains deleted passages from earlier editions, editorial notes, and various supporting documents. Users of the differently paginated Macmillan paperback edition of the eighth edition should note that all page references to the eighth edition given above must be located by using the table of correspondences appended to the paperback version.

Elements of the Economics of Industry (1892) had new editions in 1896 and 1899 and frequent reprints. The last preface is dated 1907. It is essentially an abridgement of Principles, designed to replace Economics of Industry, a book that Marshall had come to despise, quite unjustifiably. Elements contains Marshall’s fullest treatment of trades unions. Industry and Trade (1919) had new editions in 1919, 1920 and 1923 but only the first of these involved significant changes. Its three books deal with ‘Some origins of present problems of industry and trade’, ‘Dominant tendencies of business organization’, and ‘Monopolistic tendencies: their relations to public well-being’. It adopts a largely historical and comparative approach and focuses on contemporary issues. Nevertheless, it contains many passages and insights of permanent interest and warrants closer attention by economists than it has received until recently. Money Credit and Commerce (1923) had only one edition and conveys Marshall’s views on money, international trade and business fluctuations. Although blemished, it should not be dismissed.

An almost comprehensive annotated list of Marshall’s occasional writings is found in Pigou (1925, pp. 500–8), which also reprints many of the texts of these writings. Guillebaud (1961) reproduces further occasional pieces. The ‘Pure Theory’ chapters (1879a), privately printed by Sidgwick, were first published in reprint form
in 1930. A corrected and amplified version is included in Whitaker (1975). The two volumes of the latter also reproduce Marshall’s unpublished early manuscripts, mainly from the 1870s, including several manuscript chapters from the abandoned volume on foreign trade. Marshall’s important contributions to official enquiries are collected in J.M. Keynes (1926), a book that is supplemented by Groenewegen (1996).

The literature on Marshall’s life and thought is too extensive to allow for more than a highlighting of some significant contributions. The splendid memorial essay by J.M. Keynes (1924) is not to be missed, although outdated on some points, nor is the charming memoir by Marshall’s wife (M.P. Marshall 1944). Pigou (1925) includes fascinating vignettes by several of Marshall’s colleagues and friends. Guillebaud (1971) gives a nephew’s reminiscences. A major scholarly biography (Groenewegen 1995) covers Marshall’s life and thought exhaustively, while a comprehensive three-volume edition of Marshall’s correspondence (Whitaker 1996) provides much new information. Additional primary material on Marshall is provided in Raffaelli, Biagini and McWilliams-Tullberg (1995), where notes on Marshall’s 1873 lectures to women students are reproduced, and Raffaelli (1994), where Marshall’s essays on philosophical and psychological manuscripts from the late 1860s are reproduced and analyzed. See also Harrison (1963). Newspaper reports on public lectures Marshall gave during his years in Bristol are reproduced in Coase and Stigler (1969), Whitaker (1972) and Butler (1995).

Valuable overall assessments of Marshall are provided by Cannan (1924), Schumpeter (1941), Viner (1941), Shove (1942) and O’Brien (1981). Maloney (1985) studies Marshall’s involvement in the professionalization of British economics. An extensive body of detailed analysis and criticism of Marshall’s thought, mainly conducted in academic journals, continues to expand, with growing tributaries from Italy and Japan in particular. Wood (1982; 1996) assembles in eight volumes a somewhat miscellaneous collection of 239 pieces on Marshall, but standard bibliographic aids such as EconLit are recommended for a comprehensive search. The 1990 centenary of the publication of Principles produced two books of essays on Marshall (Whitaker 1990, McWilliams-Tullberg 1990) and several symposia on Marshall in economics journals. Samples of recent research can be found in Arena and Quéré (2003). On Marshall’s social and behavioural views see Parsons (1931; 1932), Whitaker (1977) and Chasse (1984). For Marshall’s views on socialism and trades unions see, respectively, McWilliams-Tullberg (1975) and Petridis (1973).

See Also

- ceteris paribus
- consumer surplus
- demand price
- external economies
- Marshall, Mary Paley

Selected works


1890b. Some aspects of competition. Presidential address to Section F of the British Association for the Advancement of Science. Reprinted in Pigou (1925).


**Bibliography**


