ECONOMIC PERFORMANCE OF ARCHITECTURAL FIRMS:
AN APPLICATION OF PRODUCTION THEORY

BY

DOUGLAS EUGENE STURGEON

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree Doctor of Philosophy in Architecture
with a minor in College Teaching
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Associate Professor Richard K. Strand, Chair
Associate Professor Michael J. Andrejasich
Associate Professor Michael T. McCulley
Associate Professor Nicholas C. Petruzzi

ABSTRACT

This dissertation explores the application of production theory to architectural firms. The basis of production theory, upon which all else depends and emerges from, is the production function. The production function is simultaneously a mathematical representation of the arrangements of inputs necessary to the production of goods and services of a firm, and a conceptualization of the underlying production process. The key benefit of understanding and utilizing production functions stems from its centrality in cost minimizing and output maximizing techniques. Through the adoption of these optimization techniques architectural firms are afforded the means of improving their economic performance. In this dissertation the focus is upon the identification of the *n*-input single-output production function best suitable for empirical study of architecture firms and subsequent use in minimizing costs and maximizing output as pathways to improving economic performance.

The maturation of production theory is largely a 20^{th} century phenomena although its antecedents date from the 18^{th} century. The historical development of production theory begins with Jacques Turgot and is advanced by such luminaries as Johann Heinreich von Thunen, Antoine Augustin Cournot, Herman Heinrich Gossen, William Stanley Jevons, John Bates Clark, and John Gustav Knut Wicksell among many others. The early historical period is brought to an end and the contemporary period born with the publication of *A Theory of Production* by Charles Cobb and Paul Douglas in which the production function bearing their names first appears. This dissertation provides a brief history of the early period and a more detailed account of the development of production functions, of the *n*-input single-output variety, as it played out in the balance of the 20^{th} century. In the process, over 50 production functions and their variations are identified and characterized. A winnowing process was developed and applied that reduced this

list to five candidate production functions. An abbreviated case study performed a statistical examination each of these forms concluding that the Cobb-Douglas and Leontief production functions presented the most viable choices for empirical study of architectural firms.

To The Father, Son and Holy Spirit

Without you I am nothing,
But with you all things are possible.

ACKNOWLEDGMENTS

I owe a great debt of gratitude to the many people who played such an instrumental role in the success of this project. I thank those family and friends who were the greatest source of encouragement in the long days and dark nights of this journey. To my sister Deb, my friends Randy, Ruth, Mia, Michael, Birge, and Regenia, many thanks. To my pastor Brett, and fellow congregants Marc, Ashlee, Joel, Elizabeth, Michael, and Tammie, you are a blessing. I owe a special thanks to the faculty and students of the School of Architecture too numerous to mention; I single out my two mentors Professors Art Kaha and Michael Andrejasich, thank you for taking an interest in me and guiding my footsteps. I am deeply indebted to a group of unsung heroes of any research endeavor, the men and women who make the library system of the University of Illinois one of the premier libraries in the world, my deepest appreciation for the support you have provided. I wish to acknowledge the vast number of economists upon whose work this dissertation is built; without their theories this project would not be possible. Finally I wish to thank Dr. Richard Strand for being my advisor and dealing with my cantankerous nature, and the other members of my dissertation committee, Dr. Nicholas Petruzzi, and Professors Michael Andrejasich and Michael McCulley, thank you very much. Without all their help this project would never have been completed. For any and all errors in this dissertation I remain solely responsible.

Table of Contents

CHAPTER 1

INTRODUCTION

> …architects have not yet found the key to the dilemma of maintaining their high standards of professional service at satisfactory levels of personal earnings and firm profitability…
>
> Case and Company, The Economics of Architectural Practice[1]

Background

By the autumn of 1966 the American Institute of Architects (AIA) had become sufficiently concerned about the business environment in which architectural firms operated to commission a study by the consulting firm of Case and Company into the nature of operating costs of member firms.[2] The passage quoted above appears in the final report under a section entitled 'Diagnosis of the Architect's Dilemma' and summarizes the effects of many of the shortcomings and difficulties in managing an architectural firm that were uncovered by the study group.[3] The 'Architect's Dilemma' remains largely unresolved nearly a half century later. The Case and Company study exposed the extent to which the architectural industry was failing to achieve financial success and laid much of the blame on the industry's inability to understand the basic costs of providing design services. The basic premise of this dissertation is that an important distinction exists between 'accounting for the cost of design services' and understanding the fundamental nature of the cost structure of the design firm as revealed through cost analysis. The former is a bookkeeping function that results in financial reports and performance parameters (e.g. billing rates, direct and indirect costs, and a range of performance ratios), the latter consists of detailed analysis of costs to determine their behavior – the why and

[1] Case and Company Inc., *The Economics of Architectural Practice* (Washington, D.C.: American Institute of Architects, 1968). 63.
[2] Ibid., 1.
[3] Ibid., 63.

how of their variation. In this introductory chapter several related and fundamental issues are addressed – has the financial health of the architectural profession changed much over the last half century, how did this situation come to exist, and what action we may now undertake to better understand the costs of design services. The genesis of this dissertation emerged from this need to find better methods of determining and understanding the fundamental cost structure of architectural design.

The evidence that the conclusions articulated in the Case Company's study remain valid still today is both anecdotal and quantitative. Anecdotally, stories regarding the poor performance of the architecture industry abound; architects tell these stories, other architects and architecture students alike listen to them – they are the fodder of cocktail parties. But are they merely apocryphal? Does the available quantitative data suggest a problem? Yes. Reports commissioned by the AIA, most notably various editions of the *Business of Architecture*, combined with census data and Bureau of Labor studies and analyses provide ample substantiation that little has changed in the financial picture of individual architects and architectural firms. The compensation of individual architects has chronically lagged below that offered their colleagues in related industries and financial performance by firms, particularly small firms, remains problematic. Although a number of factors may be cited to explain how these conditions developed and why they persist, the industry should seriously consider a new approach to understanding the cost structure of design and its impact on firm profitability. The industry must undertake a comprehensive reexamination of 'how do we know what our services cost us' and 'how can that information help us be more profitable'. This dissertation focuses on the methods of understanding the cost structure of design work, the drivers of those costs, and on means and methods the industry can employ to utilize that knowledge.

A literature review of publications prescribing management ‘best practices’ disclosed a disturbing inadequacy in their collective approach to the subject of understanding the cost of design services. The advice offered through a variety of sources, including some authored or published by the AIA, simply fail to adequately address the problem of a full and accurate accounting of the true cost of doing business. These publications do not provide architects sound advice concerning effective cost control or managerial accounting. Accordingly, architects lack the ability to accurately determine the underlying cost structure of their firms. The common thread among the ‘best practices’ approaches is the reliance upon financial reporting - balance sheets and income statements - for managerial decision-making. Financial reports are effective and reliable tools in understanding the overall financial position of the firm; however, they are not effective tools in understanding why the costs of that firm behave as they do.

The problem of not understanding the costs of production did not emerge within the architectural design industry separate and independent from similar concerns in the general business community nor did the prevailing methods of accounting for the costs of design services evolve outside of the general framework established by the accounting community. It is unreasonable to expect that the architectural profession might separately and independently develop methods of accounting for costs or for reporting financial information markedly different than those used by the general business community. In deed the opposite proved true. Architects relied upon the accounting industry for standardized accounting techniques and standard financial reporting procedures and report formats (e.g. balance sheets and income statements). Historians of the accounting profession have concluded that over reliance on these types of reports and analysis based upon them are the root causes of many faulty decisions. As outlined later in this chapter the architectural industry suffered along with many in the general business community as the accounting industry's emphasis shifted in the early part of the 20^{th} century away from managerial accounting and toward financial reporting. Architects have yet to recover from this turn of events.

The narrative history of managerial accounting and the lessons-learned derived from those experiences contained in the pages of *Relevance Lost: The Rise and Fall of Management Accounting*, by Robert S. Kaplan and H. Thomas Johnson, describes clearly the detrimental impact on managerial decision-making brought about by the switch from managerial to financial accounting in the early part of the 20^{th} century.[4] The timing of that transition negatively impacted the architecture profession. At just the precise moment the architecture profession first considered establishing standards for managerial accounting, the trend away from managerial accounting and toward financial reporting was well underway. Thus during the 1930's an opportunity to right the ship and get managerial accounting in the architectural industry moving in the right direction slipped by. However, other opportunities coming out of the field of economics were just emerging.

A review of the extensive body of literature concerning production economics reveals missed opportunities to develop cost analysis techniques in the professional service arena. Among architects the mistaken belief that architectural firms operated, or could operate, on an accounting and billing system similar to that employed by accountants, lawyers, and other professional service organizations forestalled efforts that might have led to a better understanding of the underlying cost structure of architecture firms. The use of percentages (i.e. percentage of construction cost) as the basis for a fee structure served only to commoditize architectural services. Under these circumstances the challenge to architects becomes how to

---

[4] Robert S. Kaplan and H. Thomas Johnson, *Relevance Lost: The Rise and Fall of Management Accounting* (Boston, Massachusetts: Harvard Business School Press, 1987).

complete a design and make a profit while operating within the constraint of a fixed budget.

Meanwhile, as early as the period between World Wars I and II, economists had developed methods both to analyze the movement of fixed and variable costs in accordance with levels of output for manufacturing enterprises and to calculate least cost optimization conditions. However, the architectural industry failed to consider applying these techniques to the production of designs.

Four central themes emerged from a synthesis of literature reviews in architectural practice, accounting, and economics. First, as architects sought to implement modern accounting standards in the early part of the 20th century they naturally adopted the prevailing standard, that of 'financial reporting', and in doing so missed an opportunity, along with the general business community, to continue the use of effective managerial accounting methods throughout the balance of the 20th century that existed in the latter part of the 19th and early part of the 20th century. Second, among many other culprits, ineffective cost accounting greatly contributes to the dismal financial performance of individual architects and architecture firms. Third, adoption of effective cost accounting methods is a necessary first step in any program aimed at improving the dismal financial condition of the architecture profession. And fourth, in that various methods for effective cost accounting exist, the profession ought to explore their efficacy and applicability to the daily operations of firms and should adopt the most useful. The balance of this dissertation is dedicated to the examination of one such method – economic analysis.

Exploration of economic analysis, in particular use of the production function, commands our attention in this dissertation. Although its antecedents date to the early part of the 18th century, production economics emerged alongside and as an integral part of the nascent field of neo-classical economics as it developed in the latter part of the 19th and early 20th century. In its contemporary form production economics encompasses a number of specialty fields focused upon various theories regarding production, cost, revenue, and pricing behavior affecting the decision-making of firms. Production theory is concerned with the manner in which firms employ resources (inputs) to produce the firm's products or services (outputs) with emphasis upon the efficient use of inputs with the objective of optimizing the combination of these inputs so as to minimize costs or maximize production under the condition of constrained resources, and to maximize profits.[5] Production economics provides the theoretical basis of our enquiry into the potential role of the production function as a management tool.

Objective

The objective of this dissertation is to determine which production function(s) are best suited for empirical analysis of the cost structure in architectural firms. A supporting objective is the demonstration of techniques involved in optimizing performance.

A review of the relevant literature on the economics of production reveals a rich tapestry of thoughts, ideas, and theories regarding the general field of production economics and specifically the use of production functions. The fragmented body of literature dating from the 18^{th} to the early part of the 20^{th} century reflects the disjointed nature of the early development of the broader body of neo-classical economic thought as well as that of the specialized field of production economics. Publication in 1928 of *A Theory of Production* by Charles Cobb and Paul Douglas, in which the Cobb-Douglas production function first appeared[6], in conjunction with the

[5] James L. Pappas and Eugene F. Brigham, *Managerial Economics*, Third ed. (Hinsdale, Illinois: The Dryden Press, 1979). 201.
[6] Charles W. Cobb and Paul H. Douglas, "A Theory of Production," *The American Economic Review* 18, no. 1 (1928).

publication in 1947 of *Foundations of Economic Analysis* by Paul Samuelson[7] ushered in the contemporary period of developments in production economics. Correspondingly the volume of literature about production economics steadily increased during the middle and latter part of the 20th century and continues today. One consequence of the growth in economic thought about production economics is the proliferation of mathematical forms of the *n*-input single-output production function, standing today at more than 50 forms and variations. Mindful of the desire that architects adopt the production function as a means of economic analysis of the cost of design, a method of winnowing the list to a manageable few is critical. This journey begins with a survey of the historical development of production functions, then proceeds through a summary of the nature and characteristics of production functions, an example of optimization techniques applicable to daily decision-making, a review of the contemporary development of production functions, establishment of an evaluation methodology to construct the final short-list of useful forms, and concludes with an analysis of these forms utilizing a sample of data obtained from a functioning architectural firm.

Chapter Outline

A good story needs a beginning, middle and an end. This introductory chapter provides the beginning and includes a brief history of managerial accounting and the issues that arise from our knowledge of that history, a brief comment on the financial condition and stability of the architecture profession which provides background on the gravity of the need for change, and a discussion of the concept of 'cost structure' which provides a basis and context for many of the topics covered later in this dissertation. While production economics remains the primary focus

[7] Paul A. Samuelson, *Foundations of Economic Analysis* (New York, New York: Atheneum, 1947).

of this dissertation it is not the only method of costing a production process. Ergo, a brief discussion of two other methods follows next in order for two important reasons. First, these other methods, in concert with production economics, provide insight into all the aspects of the costs incurred by architects in the design of buildings, whereas one standing alone will not. Second, these other methods deserve consideration as viable costing methodologies and research into their application should be given equal importance as that given production economics in this dissertation. Following that discussion we turn to an introduction to production economics. This chapter concludes with a comment on the structure of the balance of the dissertation. We turn now to a brief history of managerial accounting.

Relevancy Lost; Opportunities Missed

Relevance Lost: The Rise and Fall of Management Accounting, aptly chronicles the ascendancy of managerial accounting during the 19th century and its eventual decline into irrelevance during the early part of the 20th century.[8] Accounting methods and techniques developed over a period of close to one hundred years fell into disuse replaced by a new set of metrics, mostly ratios. The architecture profession would fall subject to these same forces impacting the general business community and much like the general business community has yet to fully recover.

Prior to the industrial revolution the demands on the accounting function were simple, reflecting the uncomplicated nature of business transactions of the day. These transactions occurred between owner-entrepreneurs and the market – exchanges involving raw material suppliers, laborers, and customers. Because these transactions occurred in the market the

[8] Kaplan and Johnson, *Relevance Lost: The Rise and Fall of Management Accounting*.

measures of success were easily obtained. The proprietor required only enough cash from sales to cover the expenses of labor and raw materials, with some left over for profit. Because there were no 'layers of management' sophisticated managerial accounting was unnecessary. [9] The industrial revolution would, however, change the face of business and accounting.

The industrial revolution radically changed the business environment introducing economies of scale and a complexity in business operations previously unheralded; the effect was to create new requirements of accounting systems. Businesses became more complex hierarchical enterprises consisting of many layers and a multitude of internal conversion processes within the same organization. These internal conversion processes gradually replaced external transactions previously conducted in the market. The firm made intermediate products or used supply sources embedded within their organization rather than buy them through the market in an organizational scheme referred to as vertical integration. A demand therefore arose for information about the 'price' of these internal outputs. Consequently, internal transactions became as important as market transactions. Early efforts of accounting systems focused on determining these conversion prices and calculating other parameters such as cost-per-hour or cost-per-unit. In doing so accounting systems measured the cost of labor and material and made an allocation of overhead to each process. The goal became to identify costs for intermediate and final products of the firm in order to provide benchmarks and standards for use by decision-makers. When the cost of a manufacturing process varied beyond acceptable norms managers could then take steps to identify and correct inefficiencies. Through the emerging management accounting systems owners and managers at all levels found the data necessary to support timely daily decision-making thus ensuring maximum efficiency and profitability. Early success stories

[9] Ibid., 6-7.

of this approach abound in the textile mills founded in the early part of the 19th century, the railroads formed around mid-century, and the steel mills of the second half of the century.[10]

Advances in transportation and communication during the second half of the 19th century created opportunities for large retail organizations in addition to the existing manufacturing enterprises. Retail chains such as Marshall Fields, Sears, and Woolworth began to flourish. Their measures of efficiency differed from those of the manufacturing sector. Gross margins and inventory turnover among others became the measures of effectiveness and efficiency in purchasing, pricing and retail activities. Still, very much like their manufacturing cousins these organizations focused on the efficient operation of internal processes. If each process worked efficiently then so did the whole enterprise. Management accounting systems came into being to promote efficiency in key operations of the organization.[11]

The trend toward ever more complex diversified multi-activity organizations continued into the early part of the 20th century as exemplified by the Du Pont Powder Company, itself a product of the merging of several family-owned and independently operated companies, and the newly reorganized General Motors. Faced with coordinating the diverse activities of a far flung vertically integrated manufacturing and marketing organizations the managers of Du Pont had to decide how to allocate capital. Their most enduring contributions to managerial accounting included operating budgets, capital budgets and the return of investment or ROI parameter. Faced with a diversity of product lines and an inability to coordinate corporate level functions such as marketing, purchasing and finance across product lines General Motors decentralized these functional operations and made department managers responsible for the efficiency and profitability of their segments of the company. The allocation of capital went to those whose

[10] Ibid., 7-8.
[11] Ibid., 8-9.

return on capital investment was the greatest. Thus the ROI played a key role in capital investment decisions within the corporation. By 1925 common accounting functions well known today had been fully developed; cost accounts, budgets, sales forecasts, standard costs, variance analysis, and divisional performance measures. However, at this point the evolution of management accounting systems stalled.[12]

The continued evolution toward more complex processes and organizations should have spawned a similar development in management accounting systems but failed to do so for three reasons. First, the state of data collection and analysis existing in the period from 1925 to the late 1970's made it increasingly difficult to efficiently perform analysis of critical information regarding internal operations and provide that analysis in a timely fashion to decision-makers. Before the data could be collected and analyzed the situation that data was meant to describe had changed. Management information systems, not yet widely computerized, were simply not capable of keeping pace with the speed of internal operations. Second, the decline can be attributed to the growing dominance of the use of external financial accounting statements and the collection of data to support their creation. Creditors, security exchanges, regulators, shareowners, and the government became voracious consumers of financial reports, in particular the balance sheet, income statement, and earnings' reports, and a plethora of performance measures; generally ratios of one aspect of performance to another such as price-to-earnings ratio. Given the effort required to produce these reports it may simply have been too onerous and too resource intense for companies to maintain two separate accounting systems.[13] Third, university researchers and academics failed to notice or call attention to the deteriorating position of managerial accounting. Accounting academics became enamored with the simplified

[12] Ibid., 10-12.
[13] Ibid., 12-14.

economic model of the one-product one-production-process firm. They found little of value in complicated systems of cost assignments and overhead allocation systems. The literature of the period advocated for a system of separating costs into fixed and variable elements yet were never able to explain how fixed costs were driven by production or how and why fixed or variable costs changed with levels of production. The metrics calculated for managerial decision-makers began to take second seat to performance ratios. [14]

The demise of managerial accounting lies at the feet of financial accounting and reporting. Financial reporting grew in importance as a consequence of the need on the part of capital investors for measures of the cost and profit picture of firms seeking both equity and long-term capital debt. Consequentially the accounting profession focused on creating methods, techniques and standards in support of the burgeoning field of financial reporting. This came at the expense of managerial accounting methods and techniques. As financial reporting consumed the efforts of accountants both in and out of industry, managerial accounting fell out of favor. Mangers still needed information with which to make daily decisions regarding production but increasing had to rely on financial reporting in place of managerial accounting data.

Three important consequences come out of this decline in managerial accounting. First, reporting under financial accounting proved of little aid to a manger's attempt to control costs and improve productivity. Second, such reports failed to accurately show the true costs of production. And third, the short-term focus of financial reporting, which matches period costs with period revenue, failed to recognize the long-term additions to value creation resulting from those same period costs.[15] Kaplan and Johnson conclude that for managerial accounting to regain relevancy it must perform four critical functions: provide the means to allocate costs for

---

[14] Ibid., 14-15.
[15] Ibid., 1-2.

periodic financial statements, facilitate process control, compute product costs, and support special studies.[16]

Against the backdrop of the ascendancy of financial reporting the American Institute of Architects established its first set of accounting standards in 1935.[17] The legacy of managerial accounting was not yet dead, however. The bibliography of the 1935 edition listed no fewer than nine titles dealing with the subject of cost accounting although none dealing specifically with managerial accounting. Recognizing the importance of cost accounting the 1935 edition notes:

> "The business that desires to operate at the greatest efficiency and with the least waste of time and money will ascertain the cost of each specification and of each individual sheet of drawings it produces. An accurate knowledge of these detailed costs is essential if the business is to work under a budget for a job, for a job budget wherein the drawings, sheet by sheet, are not predetermined and a cost fixed for each such sheet will be of little effect, economically. The manufacturing business that does not know the cost of the smallest product produced by it can hardly survive competition, and the architect's business does not differ from a manufacturing business in this particular. The architect who does not know whether a sheet of details or a floor plan of a known character should cost him one hundred, or three hundred, or a thousand dollars, and *who does not know why, if a sheet cost him four hundred on one job, a similar sheet containing similar information should cost five or six hundred on another job*, will not attain the financial success that an architect who does know those things exactly, and acts on that knowledge, will attain."[18]

Unfortunately the promise contained in the foregoing statement was not fully realized within the pages of the 1935 edition. For while the authors dwell in excruciating detail on the proper manner of collecting such data they failed to provide instructions regarding the means or methods of analysis and interpretation, despite their continued use of the term 'analysis'. By the 1949 publication of the *Instructions: Standardized Accounting for Architects* cost accounting had become job cost accounting and the emphasis reduced to the following statement: "The importance of Job Cost Records cannot be over emphasized. Only through their use is it possible

---

[16] Ibid., 228.
[17] American Institute of Architects, *Manual of Accounting for Architects* (Los Angeles, California: Parker, Stone & Baird Co., 1935).
[18] Ibid., 102-03. Italics added.

to determine whether or not certain classes of work and type of service are profitable."[19] The emphasis upon detailed record keeping remained while focus of analysis turned to financial ratios of direct expenses to indirect expenses among others. Their directions were clear "The **Indirect Expense Factor** is determined by dividing the total indirect expenses for the year to date by the total number of direct man hours for the year to date on all jobs."[20] Implementation of the indirect expense factor concept moved even further away from the original intent with "The Indirect Expense to be charged to each job is then computed by multiplying the number of direct man hours on the job for the year to date by the Indirect Expense Factor."[21] Any thought of computing cost of individual products such as sheets of drawings or amount of specifications was lost to the history books. Thus the ascendancy of financial reporting overtook good sound managerial accounting even within the architecture profession.

In the intervening sixty plus years AIA has changed little regarding what it recommends for proper accounting methods for the architecture profession despite numerous revisions of the instructions for accounting standards published in the same time period. Accounting methods recommended by the 14th edition of *The Architect's Handbook of professional Practice*[22] and found in *Financial Management for the Design Professional*[23] by Lowell Getz continue favoring the use of performance ratios. The use of ratios, such as the indirect expense ratio, the direct personnel expense multiplier, and utilization rates, in preparation of profit plans, annual firm budgets and job budgets are direct descendants and consequences of the use of financial reporting system metrics. Admonitions to use these metrics as the basis for daily decision-

[19] ———, *Instructions: Standard Accounting for Architects* (Washington, D.C.: American Institute of Architects, 1949). Chapter 20.
[20] Ibid., Chapter 26. Bold type original.
[21] Ibid.
[22] ———, *The Architect's Handbook of Professional Practice*, Fourteenth ed. (Washington, D.C.: John Wiley & Sons, Inc., 2008).
[23] Lowell Getz and Frank Stasiowski, *Financial Management for the Design Professional a Handbook for Architects, Engineers, and Interior Designers* (New York, New York: Whitney Library of Design, 1984).

making in production, personnel management, or budgeting are misguided. Only those cost accounting methods and techniques that seek answers to the difficult questions of why and by how much does the cost of production vary can mangers hope to have the managerial data they require in their daily decision-making.

The linkage between the decline in the use of sound managerial accounting and that of the financial performance of architectural firms has yet to be fully established. Much research lies ahead for anyone wishing to substantiate such a linkage. Nevertheless it is interesting to ponder what contribution the absence of sound managerial accounting makes to the dismal performance of many architectural firms. In the following section a small sample of the financial performance of architectural firms is drawn from the reporting of the American Institute of Architects and other sources.

Financial Health of the Architectural Industry

This section provides a snap-shot of the financial health of the architectural profession. At best it's a brief exploratory view taken through a narrow opening, but enlightening all the same. The view is disturbing; the unwholesome financial health of the industry should be a matter of great concern. A more thorough examination of the financial condition of architecture and the underlying and contributory forces and factors creating this condition commands our attention; however, limited time and resources preclude such an examination within these pages and at this time. What we can say is that the conventional wisdom that the economic fortunes of architectural firms are directly correlated to the overall national economy and to building cycles does not fully explain the fragile and disappointing economic condition of the industry. Among the other myriad causes, ranging from ineffective marketing to unprofitable compensation schemes, we must be willing to address our inability to properly determine the cost structure of design. These statements may seem overly bold and perhaps somewhat inflammatory. Before we are too quick to judge let us consider some facts regarding the overall financial condition of the industry as reported by the American Institute of Architects, as contained in the 2000 census data, from the Bureau of Labor Statistics, and finally anecdotal comments of practicing architects.

Profitability, as a measure of the surplus of revenues over expenses, affords us a basic metric to consider financial performance. 2005 is the last year for which the AIA posted profitability data as contained in *The Business of Architecture* dated 2006.[24] While 57% of firms reported double digit profits 12% reported no profit while another 17% reported profits of less than 5%.[25] For firms with more than 50% of their billings coming from the commercial and industrial sectors the numbers are 15% with no profit, 16% with a profit less than 5% - a collective 31% with a profit of 5% profit or less.[26] This is not a phenomenon of current economic conditions alone however. The report of the Case study conducted forty years earlier notes that fully 1 in 4 projects failed to generate a profit and that 1 in 12 firms failed to make a profit in 1966.[27] Profits are, however, also a function of revenue generation.

Generating profits, never an easy achievement, becomes more difficult when increases in revenue levels exhibit inconsistencies. Billings-per-employee provides one measure of revenue generation useful in comparing firms of various sizes. Revenue generation during the period 1995 to 2008 varied both by year and by size of firm. The greatest fluctuation in billings-per-employee occurred in firms of one to ten employees during the period. Billings for single person

[24] American Institute of Architects, *The Business of Architecture* (Washington, D.C.: American Institute of Architects, 2006).
[25] Ibid., 60.
[26] Ibid.
[27] Case and Company Inc., *The Economics of Architectural Practice*: 3.

firms rose from $85,000/employee initially to $96,000 at the midpoint of the period then fell to $87,000 by the end of the period. Billings for 2-4 person and 5-9 person firms saw a drop from the beginning of the period where billings were $93,000 and $87,000 per/employee respectively to $70,000 and $74,000 respectively at the midpoint then rebounding by the end of the period to $82,000 and $92,000 respectively. In contrast to small firms, the situation for firms of 10-19 and 50 or more persons exhibited more stability with a net gain of $26,000 and $33,000 per employee respectively. The biggest winner was the firm of 20-49 persons with a net gain of $46,000 per employee over the period. However, when inflation is considered the picture appears bleaker. Against a decade which saw the consumer price index rise annually by an average of 2.9 percent the lack of growth in billings for small firms is particularly alarming. For 80% of firms (those with 9 or fewer employees) the increase in billings per employee failed to keep pace with the annual rise in inflation between 1995 and 2008. Those firms of 10-19 and 50+ employees barely managed to keep pace with inflation. Only firms of 20-49 employees experienced a gain against inflation showing a 4.4% annual gain against 2.9% annual increase in inflation. In a time period of rising costs, as measured by increases in the CPI, and stagnant earnings levels the general picture of a flailing and ailing economic condition emerges.[28] If firms are struggling what is the condition of individual architects?

Wage data from both the 2000 Census and 2009 Bureau of Labor Statistics reporting indicate that individual architects are negatively impacted by this economic condition. The 2000 Census presents data showing the median annual wage of architects trailing that of electrical engineers by 7.5%, civil engineers by 10% and mechanical engineers by 12%. The Bureau of Labor Statistics reported in 2009 that the medium annual wage of architects trailing that of electrical

[28] Billings per employee and firm data is contained in the American Institute of Architects periodically published information and data about the architectural industry. The three most recent publications are the “AIA Firm Survey 2000/2002”, “The Business of Architecture” released in 2006, and “The Business of Architecture” released in 2009.

engineers by 8.5%, that of civil engineers by 3% and mechanical engineers by 2.2%.[29] This data shows the wages of individual architects trailing across the board some of the most common consultants used by architectural firms. The consultants are better off than the architects.

Finally, we turn to the anecdotal musings of architects themselves. Roger K. Lewis, a professor of architecture at the University of Maryland and noted architectural columnist for the Washington Post, comments in his widely read book *Architect?: A Candid Guide to the Profession*, "It is possible to achieve substantial wealth as an architect – and no doubt some architects pursue this as a primary personal goal – but it is improbable. Instead, most architects earn comfortable or modest livings, enjoying reasonable but *limited economic stability and prosperity*."[30] Regarding wages and compensation the general theme of his book is that architects too eagerly forego an adequate salary or sufficient compensation for a project due to the pressures of competition from other architects. Lewis further opines that architects defend and justify their actions on aesthetic grounds. If one appeals to the lofty goals of aesthetic beauty and promoting the health, welfare and safety of the general public, as common goals of architects, then surely a warm and fuzzy feeling makes up for a slightly lower paycheck. Jim Morgan in his book on managing small firms, *Management for the Small Design Firm*, tells the story of two architects at a party.[31] One asks the other what he would do with a million bucks. To which his friend responds, "Why, I guess I'd just keep on practicing until it was all gone."[32] Thomas Fisher, writing in the final issue of the journal Progressive Architecture, pens an exposé

[29] Data for the salaries is contained in the 2000 Census Report, Earnings by Occupation and Education, http://www.census.gov/hhes/www/income/data/earnings/call2usboth.html
[30] Roger K. Lewis, *Architect?: A Candid Guide to the Profession* (Cambridge, Massachusetts: The MIT Press, 1998). 1. Italics added
[31] Jim Morgan, *Management for the Small Design Firm: Handling Your Practice, Personnel, Finances, and Projects* (New York, New York: Whitney Library of Design, 1998).
[32] Ibid., 45.

on the compensation of architects titled *Who Makes What and How We Might All Make More*.[33] Fisher points out that an architect's sense of compensation includes the pure gratification of the work, as one architect related to him "Face it, architecture is a lot of fun." Fisher goes on by noting the disparity between starting salaries among architecture and engineering students. In 1995 dollars architecture students made an average of $22,125 versus $35,350 for engineering graduates.[34] Granted his data is now 15 years out of date it none-the-less demonstrates the enduring iniquity between the professions. The words of Roger Lewis echo again "Be an architect for many *other reasons* but not to get rich."[35]

One would be hard pressed to reach a definitive judgment regarding the nature of the economic condition of the architecture industry from the foregoing limited sample of data and observations. Nor, given limited time and resources, was that the intent of these few paragraphs. But it should be sufficient to suggest strongly that all is not well either. Among the many possible factors undermining a sound economic foundation we should at least consider the manner in which we understand, calculate, and express the cost of design. If we cannot, or will not, our efforts to improve profitability are forestalled before they begin.

Cost Structure

An understanding of what constitutes the 'cost structure' of an architectural firm plays a critical role in the discussions that follow yet we are hampered by the fact that no single commonly accepted or conventional definition of 'cost structure' exists. We are obliged to

[33] Thomas Fisher, "Who Makes What and How We Might All Make More," *Progressive Architecture*, no. Dec (1995).
[34] Ibid., 50.
[35] Lewis, *Architect?: A Candid Guide to the Profession*: 23.

formulate one suitable for our purposes. We begin with an examination of different perspectives about what constitutes cost.

Catherine J. Morrison Paul provides the economists perspective in her book *Cost Structure and the Measurement of Economic Performance*.[36] She employs basic production theory whereby the total cost of the firm's production is the weighted sum of all the inputs and their respective costs (wages or rents). Utilizing the production function one can ascertain the relationship between inputs and outputs, and determine how changes in output levels affect those relationships. Employing this knowledge and the mathematical power of production functions cost minimization, output maximization or profit maximization points can be calculated, and input demand schedules constructed.

Warren, Fess and Reeve, in their book *Accounting* offer us the accountant's perspective.[37] They employ the terms fixed and variable costs to describe those costs that do not vary according to output levels and those that do vary, respectively, while mixed costs are defined as exhibiting characteristic of both types, fixed and variable.[38] These concepts are employed in cost-volume-profit analysis to determine break-even points and to conduct analysis of the impacts of changes in fixed or variable costs.[39] Alternatively costs may be assigned to a cost object – defined as a particular aspect of production. In this instance costs are designated as direct or indirect according to whether the cost can be assigned directly to the cost object or not.[40] This approach is the most common in architectural firms. Another useful approach is found in absorption and variable costing with absorption costing being defined as the total of direct (labor and material)

---

[36] Catherine J. Morrison-Paul, *Cost Structure and the Measurement of Economic Performance* (Boston, Massachusetts: Kluwer Academic Publishers, 1999).
[37] Carl S. Warren, Philip E. Fess, and James M. Reeve, *Accounting*, 18th ed. (Cincinnati, Ohio: South-Western College Publishing, 1996).
[38] Ibid., 777-81.
[39] Ibid., 786-86.
[40] Ibid., 672.

costs plus indirect costs while variable costing subtracts a portion of indirect cost not assignable to a cost object (i.e. certain overhead costs).[41] We may also view costs through prime vs. conversion costs or product vs. period costs. Prime costs are composed of direct material and labor costs while conversion costs are composed of labor and overhead costs related to the item produced. Product costs are the sum of direct labor and material plus overhead costs related to the product produced while period costs are general, sales, and administrative costs.[42] The final perspective of the accountant centers on cost behavior and cost variance analysis based on standard costs. Cost behavior, defined as the manner in which a cost varies with levels of production, examines the direction and magnitude of cost changes.[43] Cost variance analysis determines direction and magnitude of cost changes against a cost standard, defined as the computed normal cost (often a statistical average).[44] Clearly the accountant is adept at slicing and dicing costs to meet the manager's decision-making requirements.

The advocates of activity-based-costing (ABC) provide a third perspective. ABC finds its antecedents in the previous referenced work of Robert Kaplan, *Relevance Lost: The Rise and Fall of Management Accounting*, and his primer on ABC titled *Cost and Effect*.[45] In describing the logic of ABC Douglas T. Hicks states the following: "The concept is simple… Costs are either assigned directly to a job, product, or service, or they are assigned to the various activities performed by the organization".[46] Those costs assigned to the organization's activities are eventually assigned to its jobs, products, or services as the cost of the various activities are associated with jobs, products or services that made them necessary." Let us take two examples

[41] Ibid., 797-80.
[42] Ibid., 674-75.
[43] Ibid., 866.
[44] Ibid., 863.
[45] Robert S. Kaplan and Robin Cooper, *Cost & Effect: Using Integrated Cost Systems to Drive Profitability and Performance* (Boston, Massachusetts: Harvard Business School Press, 1998).
[46] Douglas T. Hicks, *Activity-Based Costing: Making it Work for Small and Mid-sized Companies,* Second ed. (New York, New York: John Wiley & Sons, Inc., 1999). 50-51.

applicable to an architectural firm. The man-hours consumed directly in the production of a design are direct costs and are assigned directly to the project at hand. That part of the marketing or human resource activity associated with a particular category of building typography is assigned the indirect or support activity cost consumed by the support activity's action for that building category. In other words, if the marketing department expends resources promoting a given category, perhaps office buildings or K-12 educational facilities, then that cost is allocated to that building category and then allocated to specific projects within the category. In this manner the proponents of ABC argue, that as managers need more accurate information about the costs of processes, products, and customers than obtainable from the system used for external financial reporting, ABC meets that need by enabling indirect and support expenses to be driven, first to activities and processes, and then to products, services, and customers, thus giving managers a clearer picture of the economics of their operations.[47] This process is critical as overhead or fixed expenses are not fixed but rather 'super variable' and often discretionary thus direct costing strategies are incapable of accurately allocating these costs.[48] Moreover ABC cost systems avoid the period cost matching problem. Under GAAP (Generally Accepted Accounting Practice) period costs are matched with period revenues in the determination of profit, however, the impact of costs incurred by an activity during one period may not be felt for several future periods. Take the marketing example again. The marketing effort spread over several prior reporting periods may just now bear fruit with increasing design commissions even as the marketing expense levels off or declines in the current reporting period. ABC affords a system that looks out over multiple reporting periods to obtain a more realistic sense of the cost structure of the firm. This is possible because ABC identifies the activities and processes

[47] Kaplan and Cooper, *Cost & Effect: Using Integrated Cost Systems to Drive Profitability and Performance*: 3.
[48] Ibid.

required in the operation of the business, tracks what resources and how many resources are consumed in those activities and processes, and how much of each activity is required by each product, service or customer,[49]

In light of the foregoing we now consider the perspective of architectural firms. First, we should consider what resources are consumed in the process of producing a building design or any of its constituent parts and separate functions and/or deliverables. Here architectural firms are unique even among professional service firms. Beyond the dollars normally associated with costing strategies architects must be concerned with the consumption of time and specialized talents. Time itself consists of two components namely the time in man-hours required to complete a project or deliverable and the time-span required to produce that project or deliverable. In manufacturing circumstances production is consider instantaneous, not so with the design of building projects that may take weeks or months to complete. Architects must also consider specialized human talents or knowledge. Not every employee is capable of writing specifications, analyzing structural systems, or performing other highly specialized tasks. The architectural manager must balance each of these resources to successfully complete a design. Most importantly we should keep in mind the admonition of the 1935 standard of accounting practices mentioned earlier namely that an architectural firm should "ascertain the cost of each specification and of each individual sheet of drawings it produces.", and that "The architect who does not know whether a sheet of details or a floor plan of a known character should cost him one hundred, or three hundred, or a thousand dollars, and *who does not know why, if a sheet cost him four hundred on one job, a similar sheet containing similar information should cost five or*

[49] Ibid., 79.

six hundred on another job will not attain the financial success that an architect who does know those things exactly, and acts on that knowledge, will attain."[50]

What then constitutes the 'cost structure' of an architectural firm? First it, must recognize the various types of resources consumed in the operations of a firm, namely that the costs are monetary (dollars), time (man-hours and project completion duration), and utilization of specialized knowledge or talents as represented by uniquely skilled or trained employees. Second, it must compute the direct costs of the provision of products or services and make appropriate allocations of indirect costs associated with support activities and general overhead expenses. Third, it must recognize cost variances, their magnitude and direction, and give an understanding of the underlying cause and effect. Fourth, it must provide information useful in wide-ranging managerial decision-making contexts, including but not limited to, production optimization, man-power acquisition and retention, composition and sizing of the workforce, performance evaluation of supporting activities, project bidding and budgeting, operating budgets, and future business development. Overall, the cost structure must articulate the cost, in terms of consumption of resources, to complete projects, account for the contribution and performance of supporting activities, explain variances in economic performance and permit realistic project bidding and budgeting. Estimating the actual cost, whether in dollars, time, or human knowledge requires a robust toolbox of techniques and methods for any one technique or method may prove insufficient. In this dissertation production economics receives our prime attention but we would be remiss not to briefly explore two other methods as important tools in our toolbox and to give credence to efforts to further explore their application in architectural practice.

[50] American Institute of Architects, *Manual of Accounting for Architects*: 102-03. Italics added

Alternative Methodologies for Computing Cost Structure

Three methods of computing various aspects of the cost structure of architectural firms avail themselves. The first, production economics, is the prime focus of this dissertation. Because it is dealt with extensively elsewhere in these pages the other two methods receive our attention here. The accounting method and a statistical method are briefly developed simply to indicate their potential as computational aids and in the hope that further research into their application to architectural firms may soon be undertaken.

Accounting Analysis

Activity based costing (ABC) forms the basis of the accounting analytical methodology and was briefly introduced earlier. Robert S. Kaplan and Robin Cooper have pioneered the latest and most modern approaches to cost accounting. Their most recent publication *Cost and Effect* represents their current thinking on activity based costing and its implementation.[51] The concept of activity based costing began in the 1980's and grew out of a need to better understand how various functions and processes contributed to cost, profits, and overall efficiency in business.[52] In seeking answers to the prevailing economic questions facing architectural firm managers, ABC provides valuable insights into accurate measurement of the costs of various activities of the design production process and affords a superior methodology for assignment of indirect or overhead costs.[53]

Implementation of an ABC analysis entails a four step process involving activity definition, activity costing, allocation of costs to products, and selection of cost drivers. Essentially, organizations spend money on discrete activities which when linked in a logical

[51] Kaplan and Cooper, *Cost & Effect: Using Integrated Cost Systems to Drive Profitability and Performance*.
[52] Ibid., 3.
[53] Ibid., vii.

order or process result in the production of the company's product. ABC seeks to understand what activities are occurring, how much money or resources are spent on those activities, how these activities relate to the finished product, and how to allocate the costs of these activities to the final product.[54]

The first step in the process is developing the activity dictionary. The objective is a set of interconnected activity labels describing both activities directly associated with production and a set of indirect activity labels supporting the overall function of the firm. These activity labels are action oriented and describe a discernible result or tangible product. Generally the labels include action verbs and the desired result of that action. Develop marketing brochure, contact new customers, hire new employees are examples of such labels and are preferred over labels such as marketing, customer relations, or human resource department. Direct production activities might include produce site analysis, prepare specifications, or design foundation system. The resulting set of interrelated direct and indirect activity labels constitutes the activity dictionary.[55]

The second step in implementing ABC involves identifying the costs associated with each item in the activity dictionary. The first phase of this step typically requires allocating costs of salaries, material, and overhead expenses related directly to each activity. Overhead costs include fair share of such items as rent, utilities, equipment, etc.[56] the chart depicted below shows how the typical categorical groupings shift from expense categories to activities performed. [57]

---

[54] Ibid., 79-85.
[55] Ibid., 85-86.
[56] Ibid., 86.
[57] Ibid., 87. This is the author's reproduction of the chart appearing on page 87 of Chapter 6.

Shifting from expenses to activities alters perspective

Salaries and Fringes $313,000
Occupancy $111,000
Equipment and Technology $146,000
Materials and Supplies $30,000
Total $600,000

Activity	Salaries and Fringes	Occupancy	Equipment and Technology	Materials and Supplies	Total
Process Customer Orders	$31,000	$5,300	$12,600	$800	$49,700
Purchase Material	34,000	6,900	8,800	1,500	51,200
Schedule Production Orders	22,000	1,200	18,400	300	41,900
Move Materials	13,000	2,100	22,300	3,600	41,000
Set Up Machines	42,000	700	4,800	200	47,700
Inspect Items	19,000	13,000	19,700	800	52,500
Maintain Product Information	36,000	2,800	14,500	400	53,700
Perform Engineering Changes	49,000	3,200	26,900	2,400	110,300
Expedite Orders	14,000	900	700	500	16,100
Introduce New Products	35,000	44,000	16,100	18,700	113,800
Resolve Quality Problems	18,000	2,100	1,200	800	22,100
Total	$313,000	$111,000	$146,000	$30,000	$600,000

Figure 1, Allocation of Costs to Activities

In the second phase of step two the activities are given a set of attributes. One attribute is the position the activity holds in the hierarchy of activities within the organization. Additional attributes include but are not limited to characterization of the variability of the cost of that activity, personnel associated with the activity, and its economic performance.[58]

In step three the ultimate beneficiary of the hierarchy of activities namely the products, services and customers are identified and linked to the activities creating or supporting them. In this manner the cost of products or services can compared with the revenues they generate. Whether the supporting activities are worthy of their place in the hierarchy can be assessed.[59]

58 Ibid., 92-94.
59 Ibid., 94-95.

The final step of implementation is determining the appropriate cost driver for each activity. For some activities within the architectural firm the task, transaction, or square feet may be appropriate, while others may be tracked on some other basis. For example, the costs of some design functions vary directly with the amount of square feet being designed. For other activities such as site analysis costs are less sensitive to the amount of square feet of the design as their costs are functions of other characteristics of the task. When hiring employees to meet general operating requirements, such as a secretary/receptionist, and the cost of hiring them should be broadly allocated. However, for those hired for specific projects the cost of hiring them should be allocated to those projects. Thus the activity driver becomes a means of understanding the cost of the activity and how it should be allocated.[60]

When complete the ABC analysis provides a means of understanding the internal linkage of activities, their costs, how they relate to the finished product and customer served.

Statistical Analysis

The second analytical methodology is derived from the property valuation methodology common to the real estate industry. Property appraisers have a rich literature describing three broad systems for property valuation. His book *Real Estate Valuation: Principles and Applications* Kenneth M. Lusht describes these methods as the sales comparison, replacement cost, and income generating approaches. [61] Lusht describes the sales comparison approach as consisting of the direct sales approach, a basic statistical approach, and a multiple regression

---

[60] Ibid., 95-99.
[61] Kenneth M. Lusht, *Real Estate Valuation: Principles and Applications* (State College, PA: KML Publishing, 2001).

approach. The statistical analysis approach advocated here is a derivative of the multiple regression analysis approach from his book.[62]

The real estate industry uses multiple regression analysis in determining property values by regressing known characteristics of a group of properties against their selling prices or value. Appraisers begin with a list of characteristics describing a group of properties. These attributes are believed to have a cause and effect relationship with the value of the property. In residential property for example, the size of the house in square feet, the number of bedrooms, number of baths, and whether the house has a garage or fireplace are among a long list of possible characteristics that likely effect the selling price of a house. The relationship can be expressed in functional notation setting the selling price V as a function of the characteristics of the house (size, # of rooms, etc.) or $V = f\left(\text{characteristics of the property}\right)$. In regression notation the basic function is symbolized $V = \alpha + \beta_1 x_1 + \beta_2 x_2 + ... + \beta_n x_n + \varepsilon$, where the x_i's are the characteristics of the house, the β_i's are the coefficients of regression, α is the intercept and ε is the error term. Once standard ordinary least squares analysis is performed on the data set the standard error of the coefficient and its associated t value are examined to determine if one or more independent variables should be removed from the overall model. Researchers make the final decision as to which variables to include or exclude such that the statistical model makes sense as a mathematical representation of the real world. In some cases removal of an independent variable would make the resulting formula less convincing as an explanatory tool of what is occurring in the real world indicating that it should be retained in the regression formula.

The adaption utilized here is to substitute the cost of design, time expended, or unique resource consumed, for the house value and a list of building characteristics upon which the

---

[62] Ibid., 139-66. The discussion which follows is a synopsis of Chapter 8, Sales Comparison Using Regression Analysis.

design is developed, for the original building characteristics. Rather than number of rooms, etc., as used in the house example, the design considerations of total square feet, ratio of gross to net square feet, type of foundation or enclosure system employed, or presence of an atrium or courtyard are among a long list of possible characteristics used, depending upon the building type. The statistical analysis process remains the same as for the house example while yielding a great deal of information regarding the causes of variation in each of the architect's cost areas. This type of analysis should produce usable results at various levels of resolution ranging from the project level down to the phase, task or sub-task level. As we move from among levels of analysis – project, phase, or task – the mix of statistically valid characteristics likely changes also. Still each separate analysis may yield valuable information useful in constructing budgets, developing a bid, or in managing the project or firm. In this manner either the time consumed or cost of design may be studied and aid greatly in managing the firm.

A General Note on Methodology

This note is to answer the question "how should we proceed?" In appearance, production functions look like ordinary multivariate functions and as such it is reasonable to expect them to be revealed through multivariate regression analysis and standard model building techniques. That analysis would set the output of the architectural firm as the dependent variable with the various factors of production, mostly labor in the instance of an architectural firm, as the independent variables. However, two problems present themselves. The production function is simultaneously a conceptualization of the economic relationship between inputs and outputs in a production process and a mathematical construct. As a mathematical construct the production function must behave in an acceptable manner and observe certain restrictions, which we shall encounter in chapter 3. For example it must exhibit convex isoquants. Given that we cannot guarantee an equation developed through multiple regression analysis would behave in an acceptable manner we are left with the task of extensively proving its characteristics and modifying it if and when it fails to meet requirements. Alternatively we have the option of surveying those production functions already developed and proven by economists over the last 75 plus years. We will see in chapter 4 that over 50 plus such production functions and their variations already exist; perhaps we are better off taking advantage of them rather than reinventing the wheel. The second problem centers on the conceptualization issue. The production function should possess some power to describe how and why certain relationships exist between inputs and outputs and reflect the entrepreneurs understanding of the production process. That is it should have a basis in production theory. An extensive literature espousing production theory and the production function already exists and it seems a more logical starting point in determining the appropriate form of the production function that relying upon a pure statistical strategy. Confronted with a choice among a great many production functions we encounter a second "How do we proceed?" question.

Our direction is suggested by the work of Dhammika Sharmapala and Michael McAleer in, *Econometric Methodology and the Philosophy of Science*[63], and that of Spiro Stefanou and Kristiaan Kerstens to summarize the preceding body of work in, *Applied Production Analysis Unveiled in Open Peer Review: Introductory Remarks*.[64] Three approaches to econometric analysis are posited; traditional, instrumentalist, and falsificationist approaches. In the traditional or Cowles Commission approach the researcher accepts existing economic theory as *a priori*

[63] Dhammika Dharmapala and Michael McAleer, "Econometric Methodology and the Philosophy of Science," *Journal of Statistical Planning and Inference* 49, no. 1 (1996).
[64] Sprio E. Stefanou and Kristiaan Kerstens, "Applied Production Analysis Unveiled in Open Peer Review: Introductory Remarks," *Journal of Productivity Analysis* 30, no. 1 (2008).

true. The role of the econometrician then is that of simply estimating the value of the known parameters. No attempt is made to pass judgment on the validity of the underlying theory. The instrumentalist approach treats theory as an approximation of the real world, a false but convenient fiction. The role of the econometrician is to determine the validity of the theory through its power to predict events and outcomes. The resulting theory may still be a convenient fiction but one whose predictive powers are useful in explaining the real world. The falsificationist approach attempts to refute established theory and uses that refutation as the building block for developing new or refining existing theory. The econometrician's role is that of testing theory. Which to choose?

The approach adopted here is a combination of the traditional and instrumental. While production theory readily admits to the convenient fictional characterization of production functions, our task is the discrimination amongst numerous competing production functions rather than rendering a judgment on theory. The question then becomes "How to select from among numerous production functions?" A two prong approach is employed in this paper. In the first, utilizing criteria established by Melvyn Fuss in *Production Economics: A Dual Approach to Theory and Applications*,[65] the number of candidate production functions is winnowed to a manageable few suitable for empirical study of architectural firms. The final selection is likely to be highly firm specific based upon their unique business circumstances and process model. In chapter 6 an analysis of one mid-west firm is presented in which the goodness of fit and standard error are the main characteristics of the production function under evaluation.

---

[65] Melvyn Fuss and Daniel McFadden, *Production Economics: A Dual Approach to Theory and Applications, Vol I and II* (New York, New York: North-Holland Publishing Company, 1978).

Plan of This Dissertation

Here in chapter one the background to this dissertation was established along with the objective of the research. The concept of the cost structure of architectural design was introduced and mention was made to two competing alternative methodologies for establishing the exact nature of that cost structure. The focus none-the-less is upon production economics and the *n*-input single-output production function.

Chapter 2 is devoted to an early history of production economics and through that history an introduction to important concepts developed prior to the early portion of the 20th century is presented. Chapter 3 presents the essential basics of production economics as developed throughout the 20th century with emphasis on the production function and its use in *n*-input single-output scenarios. Chapter 4 presents the development of the production function in contemporary times enumerating and characterizing more than 50 functions and their variations. Chapter 5 develops a winnowing methodology. By employing this methodology this large number of candidate production functions is reduced to a small number suitable for statistical analysis. Chapter six presents a short case study utilizing data from a mid-west firm to demonstrate how the final selection may be made. The final chapter, chapter 7, provides a summary of this dissertation, presents the conclusions drawn from it, and discusses various avenues for additional research.

CHAPTER 2

HISTORY OF PRODUCTION THEORY

Discovery consists of looking at the same thing as
everyone else and thinking something different.
Albert Szent-Györgyi de Nagyrápolt[66]

Introduction

The economic literature regarding production functions commonly, yet unfortunately, overlooks, ignores or disregards the contributions of numerous early economists, and those non-economists, who, over a period of nearly two centuries, brought the development of production economics to the threshold of the modern era. The purpose of this section is give credit where it is due, to explain the origins of the concepts utilized later in this paper as a way of promoting a more thorough understanding of this material, and establish a historical context for what follows in chapters 3 and 4. In this regard we trace the work of Jacques Turgot in the mid to late 18th century up to the works of Charles Cobb & Paul Douglas in 1928, and that of Paul Samuelson in 1947. Several excellent accounts provide details of this period in the development of economic thought, including production economics but extending beyond it also, and I do not propose to replicate in its entirety that history here.[67] However, a short treatment of the salient points pertinent to production economics is appropriate.

Background

Following the end of World War II the 21st century witnessed a grand profusion of ideas as the field of production economics blossomed. Following the scheme set out by Jacob Oser

[66] Nobel Media AB, http://nobelprize.org/nobel_prizes/medicine/laureates/. The Nobel Prize in Physiology or Medicine 1937 was awarded to Albert Szent-Györgyi *"for his discoveries in connection with the biological combustion processes, with special reference to vitamin C and the catalysis of fumaric acid"*.
[67] See Bibliography for more detailed accounts of relevant historical events.

and Stanley L. Brue in *The Evolution of Economic Thought* the preceding historical period is divided into thirds.[68] The first period is comprised of the works of the forerunners of the marginalist school, Jacques Turgot (1727-1781), Johann Heinreich von Thunen (1783-1850), Antoine Augustin Cournot (1801-1877), and Herman Heinrich Gossen (1810-1858). The second period highlights the work of the marginalist school, William Stanley Jevons (1835-1882), Francis Y. Edgeworth (1845-1926), Leon Walras (1834-1910), and John Bates Clark (1847-1938). The concluding period highlights the work of the early neoclassical economists, John Gustav Knut Wicksell (1851-1926), and Alfred Marshall (1842-1924). The modern era overlaps the early developmental period as its beginning is marked by the publication of *A Theory of Production* by Charles Cobb and Paul Douglas in which the production function bearing their names first appears.[69] The modern era, however, would not take off until after the hostilities of World War II ended and the publication of Samuelson's *Foundations of Economic Analysis*.[70] Today we enjoy a rich body of literature on production economics and the field itself has grown to embrace more than just theories of production functions. It is to the early pioneers that a largely unpaid debt is owed. The following short summarization of that work pays homage to their efforts.

Nascent Neoclassical Economics

The history of the emergence of production theory as a gradual coalescence of the thoughts and deliberations of great minds of the 18th, 19th and early 20th century plays out in the shadow of the broader ascension of neoclassical economics over classical political economy.

[68] Jacob Oser and Stanley L. Brue, *The Evolution of Economic Thought*, Fourth ed. (SanDiego, California: Harcourt Brace Jovanovich, Inc, 1988).
[69] Cobb and Douglas, "A Theory of Production."
[70] Samuelson, *Foundations of Economic Analysis*.

The publication of the *Wealth of Nations* [71]in 1776 by Adam Smith (1723-1790) represents the pinnacle of classical political economic thought and held sway throughout most of the 18th and 19th century, yet the landscape of economics was changing.[72] The early stages of the industrial revolution had a transformative effect on national economies. The focus shifted from what the king acquired in wealth to what the nation could produce. Classical economists such as Smith were moral philosophers attempting to explain the workings of this new economic system in which men lived through observation of facts and deductions regarding the underlying cause and effect relationships. They concerned themselves with the production, distribution, exchange and consumption of national wealth with focus on how and why it changed over time – they thus constituted the first group of macroeconomists. What began with Jacques Turgot in the later part of the 18th century was bearing fruit and beginning to emerge as a coherent body of theories throughout the later part of the 19th century and into the 20th century. Nascent neoclassical thinkers turned their focus away from the larger questions of wealth accumulation over time and toward systems that explained consumer and producer behavior. The key question became how to allocate scarce resources through a market system so as to maximize satisfaction of consumers, efficient allocation of resources, and profits of producers. The neoclassical shift was from the aggregate of national systems to the acts and actions of individuals within the national economy. The goal was to find the equilibrium point or optimum allocation of resources. The methods grew evermore mathematical in nature and moved away from the moral philosophy of classical political economy. In this way the neoclassical economist became the first of the micro

[71] Adam Smith, *An Inquiry into the Nature and Causes of the Wealth of Nations* (London, England: Strahan and Cadell, 1776).
[72] Jurg Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980* (Baltimore, Maryland: The Johns Hopkins University Press, 1990). 60-72.

economists.[73] Before neoclassical economists were known by that term, before the marginalists that preceded them and gave us mathematical economics came the first to sow the seeds of eventual neoclassical thought the forerunners of the marginalist school.

Forerunners of the Marginalist School

Jacques Turgot (1727-1781) born in Paris, he aspired to enter the clergy but entered instead into government service where he served as chief administrator of Limoges under Louis XV and minister of finance under Louis XVI.[74] Turgot was the first to articulate the 'law of returns' what we today refer to as the 'law of diminishing returns' and thus laid the foundation for the concept of marginal productivity. In his '*Observations on a Paper by Saint-Peravy*' (1767) Turgot describes the effect on agricultural production of increasing preparation of the field (e.g. tilling and fertilizing).[75] In effect he states that as increases in the preparation of the field occur the output of the field initially increases at an increasing rate until some level of preparation takes place such that output continues to increase but at a decreasing rate, then finally total output actually decreases with additional preparation. Turgot's rudimentary explanation sufficed until expanded upon by Thunen and the advent of production functions.[76] Turgot's life as a government administrator afforded him little time for research and original writing thus he never expressed his concept in a workable mathematical formulation.[77] The world of economic thought would wait a half a century for the work of Antoine Cournot to remedy the situation.

---

[73] William Stanley Jevons, *The Theory of Political Economy* (Harmondsworth, Middlesex, England: Penguin Books Ltd., 1871). This paragraph is a summary of that provided in the Introduction by R. D. Collison Black on pages 8-9
[74] Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980*: 73.
[75] Ibid., 75.
[76] Ibid., 76.
[77] Ibid.

The man most deserving of the title 'father of production economics' and one of the leading theorists of the pre-marginalist period was Johann Heinreich von Thunen (1783-1850).[78] Thunen's theories and observations arose from his work on his estate near Mecklenburg in Germany. Although born on his father's estate he grew up living in a small coastal town following his father's death. Facing the prospect of managing one of his father's two estates he undertook a farming apprenticeship. Disliking the hard manual labor he turned to a study of agriculture, mathematics and economics at a college near Hamburg then at Celle and later at the University of Gottingen. While at Celle he studied under Albrecht Thaer the leading proponent of rational agriculture which focused upon soil characteristics and fertility, and upon the proper method for crop rotation as the means of achieving maximum efficiency in agricultural production. Thunen was critical of this approach as it lack an economic explanation of efficiency. Thunen's conclusion held that relative prices of the inputs to agricultural production were the key to determining the optimal production mix.[79] Here then we see the seeds of what will eventually become the strategies of cost minimization. Based upon his observation of agricultural outputs he was the first to give algebraic expression to the production function. Thunen's production function found in Volume II of his work *The Isolated State*, is $p = hq^n$; where $p = \text{output per worker}$, $h = \text{a constant of production}$, $q = \text{capital per worker}$, and $n = \text{a fraction between zero and one}$.[80] Since total output capital 'P' is the product of production per worker times the number of workers or pl and q is capital 'C' per worker or C/L we can

[78] Ibid., 164-65.
[79] Ibid.
[80] Johann Heinrich Von Thunen, *Von Thunen's Isolated State*, trans. Carla M. Wartenberg (Oxford, England: Pergamon Press, 1966).

rewrite Thunen's equation as $P = pL = hL(C/L)^n = hL^{1-n}C^n$.[81] This is the Cobb-Douglas production function written almost a century before the now famous article which gives it the name by which it is most widely known. Thunen's other important contribution came in the concept of marginal productivity. Although he failed to use the term 'marginal productivity', he nonetheless described the phenomena of optimizing the output of several inputs by stating that optimization occurs "where the incremental output of the last worker is absorbed by the wage that he receives".[82] By extension this implies that the marginal product produced by the last increment of all inputs must equal the ratio of the wages of the input to the prices of those outputs. The precise calculus required to calculate the least cost point would have to wait for Herman Amstein whom we encounter shortly.[83] Through his contribution of the earliest production function and the concept of marginal productivity Thunen's place as the father of production economics is secure.

Antoine Augustin Cournot (1801-1877) was born in the small town of Gray in France.[84] In 1821 he was admitted to the Ecole Normale Superieure and began his initial study of mathematics but soon the school closed and he transferred to the Sorbonne where he obtained his degree in mathematics in 1823.[85] In the early years of Cournot's career he served as the private secretary of a French Marshal assisting in writing the memoirs of his campaigns of 1812-1813 and tutoring the his son. In 1834 he became a mathematics professor at Lyons, and then later served as the Rector of the Academy at Grenoble, Inspecteur General des Etudes, and Rector of

[81] Thomas M. Humphrey, "Algebraic Production Functions and Their Uses Before Cobb-Douglas," *Economic Quarterly* 83, no. 1 (1997): 63-64.
[82] Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980*: 170.
[83] Humphrey, "Algebraic Production Functions and Their Uses Before Cobb-Douglas," 66-68.
[84] Takashi Negishi, *History of Economic Theory* (Amsterdam, The Netherlands: Elsevier Science Publishers B.V., 1989). 243.
[85] Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980*: 176.

the Academy at Dijon.[86] During his career he published several works on differential calculus, algebra, geometry and probability.[87] When his attention turned to economics he used differential calculus to describe the various workings of monopoly and duopoly. In his most noteworthy publication, *Researches into the Mathematical Principles of the Theory of Wealth* (1838) he utilized calculus to fully describe what we today would call marginal revenue and marginal cost.[88] Jurg Niehans notes "Inasmuch as marginalist economics can be interpreted as a rewriting of classical theory in terms of calculus, Cournot provided the keynote for a century of economic theory."[89] Thus Cournot became the first economist to apply calculus to economic analysis and in doing so introduced mathematical economics to the world.[90]

Herman Heinrich Gossen (1810-1858), born near Cologne, Germany studied law and public administration at the Universities of Bonn and Berlin but his government service was unrewarding. He retired from public life in 1847 and devoted his time to writing and publishing *The Laws of Human Relations and the Rules of Human Action Derived Therefrom*,[91] in 1854.[92] Gossen provides an indirect contribution to production economics through his depiction of the equilibrium point of marginal utility. This equilibrium point is described as that point where the last unit of money spent on one 'good' gives the same satisfaction as the last unit of money spent on another 'good'. Expressed symbolically that is when $MU_x / P_x = MU_y / P_y$; where MU_x and MU_y are the respective marginal utility of commodities *x* and *y*; P_x and P_y are their respective

[86] Negishi, *History of Economic Theory*: 241.
[87] Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980*: 177.
[88] Oser and Brue, *The Evolution of Economic Thought*: 216.
[89] Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980*: 177.
[90] Oser and Brue, *The Evolution of Economic Thought*: 216.
[91] Hermann Heinrich Gossen, *The Laws of Human Relations and the Rules of Human Action Derived Therefrom*, trans. Rudopph C. Blitz (Cambridge, Massachusetts: The MIT Press, 1983).
[92] Negishi, *History of Economic Theory*: 319-20.

prices.[93] This formulation anticipates the condition in cost minimization where the marginal productivity of one input divided by its price must equal the marginal productivity of other inputs divided by their respective prices which we will encounter implicitly in the work of Walrus and explicitly in the work of Samuelson.

Marginalist School

William Stanley Jevons (1835-1882) was born in Liverpool, England and educated at University College, London. His contribution comes from *The Theory of Political Economy* first published in 1871.[94] Previous economists contended that the value of a 'good' lies in the labor required in its production; Jevons held that value depended entirely upon utility.[95] Jevons advanced the concept of marginal utility through his theory of the law of diminishing marginal utility. He claimed that while it is not possible to directly measure utility a single individual can compare marginal utilities of several goods. If a person receives satisfaction from consumption of a given 'good' total satisfaction grows as consumption increases. However, as total satisfaction grows each additional unit of consumption produces a smaller increase in total satisfaction. The change in total consumption from one level to the next is marginal utility, or as Jevons called it, the final utility. This diminishing marginal utility concept clearly antecedes what will become the law of diminishing returns.

Francis Y. Edgeworth (1845-1926), born in Edgeworthstown, Ireland studied at both Trinity College in Dublin, and Oxford University in England. His contributions to production economics are found in *Mathematical Psychics: An Essay on the Application of Mathematics to*

[93] Oser and Brue, *The Evolution of Economic Thought*: 224-26.
[94] Jevons, *The Theory of Political Economy*.
[95] Ibid., 77.

the Moral Sciences published in 1881.[96] In Edgeworth's first contribution he introduced the idea that 'curves of indifference', what we today call indifference curves, represent a set of points on a graph yielding equal value or utility between various combinations of goods.[97] Today we draw these curves as convex to the origin. Any point along the curve, while presenting different combinations of goods – for example 2 units of good 'A' and 4 unit of 'B' or 1 unit of 'A' and 5 units of 'B' – that all points along the curve have the same utility. The concepts of isocost and isoquant curves are direct descendants of Edgeworth's indifference curves. In his second contribution Edgeworth was the first to explicitly distinguish the differences between average and marginal products and provide us with the graphs of the curves of total, average and marginal production and thus show their interrelationships in production functions characterized by variable proportions of inputs.[98]

John Bates Clark (1847-1938) represents America's participation in the marginalist revolution. Born in Rhode Island he studied at Amherst and abroad in Germany. His salient publication was *The Distribution of Wealth* published in 1899.[99] He is credited with inventing the term *marginal productivity* and providing the best explanation of the marginal productivity theory of distribution based upon the law of diminishing returns as applied to all factors of production. Up to this point the law of diminishing returns had only been applied to agricultural settings, Clark generalized this idea to include all factors and in all manufacturing circumstances. Clark explained that if all inputs are held constant save one and that additional units of that variable input are added to total output that ultimately average and marginal productivity fall even though total output may continue to increase. This occurs not because the quality of the

---

[96] Negishi, *History of Economic Theory*: 330-31.
[97] Oser and Brue, *The Evolution of Economic Thought*: 252-53.
[98] Ibid., 257-59.
[99] John Bates Clark, *The Distribution of Wealth: A Theory of Wages, Interest and Profits* (New York: Macmillan, 1899).

variable input declines but as a function of the overuse of the fixed inputs by the growing amounts of the variable input.[100] By implication if each input is paid a price/wage equivalent to its marginal product then the maximum amount of output would result for any given level or combination of inputs, a condition we know today as the maximum production for a given minimum cost. This in turn implies that production is constant in terms of returns; if inputs are doubled then output doubles. A debate soon erupted, the consequence of which was a clearer understanding of returns to scale. When returns were proportional to increases in inputs then constant returns to scale were said to exist, but if an increase in inputs resulted in greater than proportional increase in production then increasing returns to scale existed and correspondingly if an increase in input resulted in a smaller proportional increase in output then decreasing returns to scale existed.[101] Knut Wicksell would later extend this concept one more level of complexity.

Leon Walras (1834-1910) born in Evreux, France and educated at the Ecole des Mines is considered a leading member of the Marginalist School. His principle contribution of a theory of general equilibrium comes to us from *Elements of Pure Economics*[102] published originally in 1874.[103] In this theory Walras extended the partial equilibrium model of one or two commodities to a general theory of equilibrium of *n*-commodities. This theory consisted of a framework of basic price and output interrelationships of an entire economy which included both commodities and factors of production. Its purpose was to demonstrate mathematically the linkage between prices and quantities of goods produced and how an adjustment or change in one generates a new equilibrium via changes in the prices and quantities of all other goods. What is of interest in this

---

[100] Oser and Brue, *The Evolution of Economic Thought*: 260-61.
[101] Ibid., 267.
[102] Leon Walras, *Elements of Pure Economics*, trans. Willian Jaffe (The American Economic Association and The Royal Economic Society by Richard D. Irwin, Inc., 1926).
[103] Oser and Brue, *The Evolution of Economic Thought*: 340.

discussion is that the mathematical model led Walras to consider the coefficients of the factors of production to be fixed. It is this analysis around which Wassily Leontief would later generate his fixed coefficient production function bearing his name.[104] Walras would later consider the case of variable coefficients of the factors of production but his limited command of advanced mathematics proved a great hindrance in the development of a general theory of marginal productivity. In correspondence with Hermann Amstein he posed the problem of solving the set of simultaneous questions contained within his general equilibrium theory and received back from Amstein a brief tutorial in the application of the LaGrange multiplier and use of partial derivatives to obtain a minimum or maximum solution. Both gentlemen had, however, missed the essential but missing element namely a statement of quantity of production necessary to complete a statement of marginal productivity theory.[105]

Neoclassical School

John Gustav Knut Wicksell (1851-1926) born in Stockholm, Sweden and studied mathematics, physics, economics, and law at the University of Uppsala.[106] His major contributions to production theory concern the exhaustion problem and an observation on returns to scale. Wicksell along with many of his colleagues were concerned with the problem of exhausting resources in production. The problem has two faces, one, how are we assured that maximum production is obtained from a given set of resources, and two, how do we establish the optimum set of resources to achieve a given output level. This follows from the general

[104] Ibid., 341-46.
[105] William J. Baumol and Stephen M. Goldfeld, "Precursors in Mathematical Economics: An Anthology," (London, England: The London School of Economics and Political Science (University of London), 1968). As found in the selection titled “Walras, Amstein and Marginal Productivity Theory” with credit to Professor Willian Jaffe, pages 309-312
[106] Oser and Brue, *The Evolution of Economic Thought*: 296.

equilibrium theory of Walras where he successfully showed a system of production of several products to be in balance yet had not satisfactorily arrived at how to coordinate the use of resources, what today we call optimal allocation of resources. Wicksell observed that "...we must approach the subject from the standpoint of the differential calculus." In this regard he stated "If the total output of production is interpreted as a real function of the cooperating factors, …then efficiency clearly requires that each factor be used to such an extent that the loss of a small portion of it reduces the resulting output by just so much as the share of output going to that portion. …that the output shares of the various of the various productive factors must be proportional to the partial derivatives of the said function with respect to the factor in question as a variable."[107] In light of Walras theory of general equilibrium and Amstein's application of the LaGrange multiplier we see that we are only missing the objective equation of minimizing cost to complete the necessary tools to perform cost minimization analysis. In Samuelson's codification in *Foundations of Economic Analysis* we find the complete cost minimization calculations. In his second contribution Wicksell moved beyond the exhaustion question and cost minimization issues to observe that a typical firm would likely experience variable returns to scale. Early in the process of adding additional resources the increases in production would demonstrate increasing returns to scale while at some future point in adding resources the production would exhibit constant returns to scale and later still with more resources would experience decreasing returns to scale.[108]

Alfred Marshall (1842-1924) born in London, England was educated at Cambridge where he studied mathematics, physics, and later, economics. Ultimately he rebelled against his tyrannical father, who in Alfred's youth is said to have overworked him at his studies, forbid his

---

[107] Niehans, *A History of Economic Theory: Classic Contributions, 1720-1980*: 251.
[108] Oser and Brue, *The Evolution of Economic Thought*: 267.

playing chess as it was a waste of time and greatly discouraged his interest in mathematics, to become one of the most acclaimed economists of his time and accomplished mathematician. Marshall's greatest contribution to economics came in his publication *Principles of Economics* first published in 1890[109] in which he codified the principles of economics as embodied in the field of neoclassical economics. Subjects covered ranged from marginal utility to taxes and subsidies, consumer choice, distribution of wealth, production costs, and, to the bane of every student of introductory economics, the laws of supply and demand. His discourse on production theory provide the most thorough treatment of production economics to date but given his predisposition toward mathematics it is devoid of nearly all useful mathematical formulations. He attitude, expressed in his own words, indicates why this was so.

> "(I had) a growing feeling in the later years of my work at the subject that a good mathematical theorem dealing with economic hypotheses was very unlikely to be good economics: and I went more and more on the rules-
> (1) Use mathematics as a shorthand language, rather than as an engine of inquiry.
> (2) Keep to them till you have done.
> (3) Translate into English.
> (4) Then illustrate by examples that are important in real life.
> (5) Burn the mathematics.
> (6) If you can't succeed in (4), burn (3). This last I did often.[110]

His disdain for the application of mathematical formulas in economic discourse precluded him from developing the techniques of optimization revealed by Amstein they would not reappear until Samuelson. Marshall's contribution comes then in the collection of various definitions and concepts in one place, the *Principles of Economics*, in which he provides greater exposition and discussion but no unique contribution beyond the collection itself.[111]

---

[109] Alfred Marshall, *Principles of Economics*, vol. 1 (Cambridge, England: Cambridge University Press, 1890). Currently available edition is ———, *Principles of Economics*, Reprint of the Eighth ed. (Amherst, New York: Prometheus Books, 1997).
[110] Oser and Brue, *The Evolution of Economic Thought*: 273. The original quotation is from the *Memorials of Alfred Marshall* edited by A. C. Pigou (London: Macmillan, 1925), 427
[111] Ibid., 271-92.

Production Economics

Neoclassical economics emerged as the dominant body of economic theories with the publication of *Principles of Economics* by Alfred Marshall but the specialized field of production economics would not fully materialize until Paul Samuelson's book *Foundations of Economic Analysis*. As with all fields, in the social and physical sciences new advances in theory, procedures, and application occur continuously. Chapter 3 presents a summary of production economic analysis, with focus solely upon the *n*-input single-output production function. Yet the field of production economics has not been and is not static. What began with production functions expanded quickly to considerations of cost, revenue and profit functions leading to the theory of duality as espoused by Ronald Shepard in 1953.[112] To handle such difficult modeling problems as the passage of time Loftsguard and Heady utilized dynamic linear programming in 1959.[113] In a challenge of the assumption that production functions represent the best technological arrangements of the factors of production economists are exploring frontier production functions[114], data envelope analysis[115], and stochastic frontier analysis.[116] Work abounds attempting to develop new applications and to apply current analysis to new areas in both the macro-economic and micro-economic world in which we live. Mindful that the field of production economic has grown and continues to grow, it remains necessary to limit the scope of

---

112 Ronald W. Shephard, *Cost and Production Functions* (Princeton, New Jersey: Princetion University Press, 1953).
113 Laurel D. Loftsguard and Earl O. Heady, "Application of Dynamic Programming Models for Optimum Farm and Home Plans," *Journal of Farm Economics* 41(1959).
114 Chauncey T. K. Ching and John F. Yanagida, *Production Economics: Mathematical Development and Applications* (New Brunswick, New Jersey: Transaction Books, 1985).
115 M. J. Farrell, "The Measurement of Productive Efficiency," *Journal of the Royal Statistical Society* 120(1957).
116 D. J. Aigner, C. A. K. Lovell, and P Schmidt, "Formulation and Estimation of Stochastic Frontier Production Functions," *Journal of Econometrics* 6(1977).

this extant exploration to the application of production functions of the *n*-input single-output model to practicing architectural firms.

Overview of Applied Production Economics

The rise of neoclassical economics transformed the thinking and focus of economists. While still concerned with many of the same issues with which classical political economists struggled, neoclassical economists turned much of their attention upon the actions of individual consumers, giving rise to consumer theory, and that of individual businesses or firms, giving rise to production theory. Concurrent with this increased focus on individual consumers and the firm, their analysis became much more mathematical with calculus and marginal productivity analysis taking center stage. The proponents of production theory sought then to use these new tools to explain the behavior of entrepreneurs. Behavior in this context is the manifestation of the various decisions made by managers as they go about the task of producing their various goods and services. In studying the underlying economics the scope of production economics expanded to encompasses cost, revenue and profit functions, the issue of duality, and investigations into efficiency frontiers. Production function theory extended its reach to embrace not only single product production processes but also joint products from a single process, and multiple products from multiple inputs sets. What began as a relatively restricted field of study has grown vastly more complex and broad in scope. The balance of this dissertation discusses only a single product production process utilizing *n*-inputs. The examples utilized here build from a single input model to a model of two inputs and finally progresses to the more applicable *n*-input model. In the exposition that follows the essence of the firm and its production process, and that of marginal productivity theory share center stage. Upon this foundation other concepts of production theory and applied production economics depend.

Managing a firm involves a series of important decisions many of which pose dire economic consequences to the firm. Production economics is a decision science that studies the fundamental aspects of the economic dimension of these decision-making processes and provides a means of rationalizing decision-making. Such decisions or choices would not be necessary if we lived in a world of plenty. Scarcity arises when the competition for resources from alternative uses creates a demand that exceeds the available supply. Allocation of scarce resources against relatively unlimited wants and needs sits at the core of all economic studies. However, some resources are abundant and therefore lack economic value and do not play in the decision-making process. Their exclusion does not mean that they are unnecessary in the production process nor does it invalidate economic analysis. It simply means that the availability of those resources have no impact on economic decision-making. Many of the factors of production, however, are in scarce supply and because of this scarcity decisions regarding their allocation are worthy of economic analysis.[117]

[117] John P. Doll and Frank Orazem, *Production Economics: Theory with Applications* (New York, New York: Joh Wiley & Sons, 1984).

CHAPTER 3

PRODUCTION THEORY

> Facts which at first seem improbable will, even on scant explanation, drop the cloak which has hidden them and stand forth in naked and simple beauty.
>
> Galileo Galilei (1564-1642)[118]

Introduction

This chapter presents a brief statement of production theory focusing upon the *n*-input single-output production process. Although the body of production economics theory is much broader than merely the *n*-input single-output model discussed herein, its selection as the focus of this effort is not arbitrary. The two alternatives, the single-input single-output model and the *n*-input *m*-output model, were not selected as they are either not appropriate, as in the case of the single-input single-output model, or more complex than can be supported by a preliminary exploration into the application of production functions to architectural firms such as the present effort. As noted elsewhere in this dissertation such subjects as duality theory, frontier analysis, and non-parametric production functions also are beyond the scope of this effort. By the study of production functions and production theories relating to the n-input single-output model we may discern their application to architectural enterprises. Thus we may find ways to better understand economic efficiency in architectural firms and begin constructing the concept of their underlying cost structure. Subsequent research into areas such as the *n*-input *m*-output model and duality theory, among others, may expand further our understanding of the cost structure of architectural firms. Such efforts must await more time and resources than can be devoted here.

[118] John Bartlett, *Familiar Quotations: A Collection of Passages, Phrases and Proverbs Traced to Their Sources in Ancient and Modern Literature* (Boston, Massachusetts: Little, Brown and Company, 1980). 183.

Chapter two began this discourse by introducing a number of key concepts of the theory of production and by placing them in their historical context to form an introduction to this important field of economics. The present chapter extends and expands that discourse with a four-fold objective. An exposition of the theory of production as it relates to the *n*-input single-output production scenario is the primary objective of this chapter, as previously noted. Demonstrating that with only minor modifications of the definitions and principles that production economics may be equally applied to the general field of the professional service firm and specifically to the architectural firm stands as the second objective. As noted in the introductory chapter the overall objective of this dissertation is the identification of the production function(s) appropriate for detailed research into the production dynamics of the architectural firm. As such establishing a foundation in production economics sufficient to support the detailed discussion of production functions developed in the latter two-thirds of the 20th century, as presented and discussed in chapter 4, constitutes the third objective. Readers not familiar with production theory should find this chapter essential reading and a handy reference when reading the next chapter. Bridging the divide between pure theory and practical application constitutes the fourth objective of this chapter. Therein the reader may find the best rationale for why production theory bears relevancy to architectural firms and their managers.

All businesses share a common defining characteristic; namely, that they transform or convert scarce resources (inputs) into the products and services (outputs) they provide to other firms and consumers through what is commonly referred to as the production process. Production processes vary greatly ranging from those that mine or harvest natural resources (forestry, mining and agriculture), to those producing intermediate or final products (manufacturing), and extending to those providing a range of services. Most employ multiple resources in creating multiple products in a process which production economists label the *n*-input *m*-output model. Other production processes utilize only a single input to produce a single product; the single-input single-output model. A third category consists of production processes employing multiple inputs to create a single product; the *n*-input single-output model; the model upon which this chapter focuses.

The selection of the *n*-input single-output model is not made arbitrarily. Many business operations are complex undertakings utilizing many and various inputs to produce a range or variety of products, good and services; architectural firms are no exception. Often, however, even the most complex operation can be disaggregated into constituent parts and those parts analyzed separately. Obviously there exits, in the context of the architectural firm, potential benefit in studying the *n*-input *m*-output model where all the inputs of the firm are used to produce more than one output. The choice made here is to study the simpler case while reserving the more complex analysis for future research.

This chapter spotlights the basics of *n*-input single-output production theory as developed by early neoclassical economists and refined in concept and exposition throughout the 20th and into the 21st century. Marginal productivity remains the central concept in production theory just as its development played a central role in the early history of production theory. Now, however, its story is encapsulated within a more comprehensive treatment of production economics. The section detailing the theory of production begins with an exploration of the production function, its definition, nature, characteristics and constraints. Next follows an examination of marginal productivity, marginal technical rate of substitution, production surfaces, isoquants, isoclines, ridgelines; returns to scale and size; factor interdependence, separability and elasticity; stages of production, elasticity and coefficient of production; and issues of convexity and concavity. However, before launching fully into the theoretical aspects of production theory we should pause and consider the nature of the firm and establish the characteristics of the production process to form a suitable foundation and bridge to the balance of this chapter.

The first section of the chapter therefore sets out the basics of the theory of the firm, first developed by the eminent economist Ronald Coase. The second section explores in detail the production process; first by defining the concept then proceeding to an examination of the nature of inputs and outputs, the role of technology, various input-output scenarios, and finally a review of the Thompson classification scheme for production processes. Based on this foundation an exposition of the theory of production for the *n*-input single-output scenario then consumes the preponderance of this chapter. The final section discusses optimization, the practical application of production theory.

Theory of the Firm

The 21st century understanding of the nature, characteristics and properties of the firm dates from the 1937 publication of the article *The Nature of the Firm* by Ronald H. Coase.[119] Dissatisfied with prevailing explanations for the existence of firms Coase reasoned that their existence was due to the ability of entrepreneurs to obtain and organize the various factors of production more cheaply and effectively than those same input factors could be obtained from or organized by the market. Following Marshall's lead neoclassical economists considered the market price of goods and services to be determined by the forces of supply and demand coordinated through the price mechanism. They argued that the price mechanism was a

[119]Ronald H. Coase, "The Nature of the Firm," *Economica* 4, no. 16 (1937). For a lifetime's work in economics and in recognition of his contribution to the understanding of the theory of the firm Professor Coase was awarded the Noble Prize in Economics in 1991. (Nobel Media AB.)

sufficient force in the directing or coordinating the distribution of resources. Coase disagreed arguing,

> …that the operation of a market costs something and that, by forming an organization and allowing some authority (an entrepreneur) to direct the resources, certain marketing costs are saved. The entrepreneur has to carry out his function at less cost, taking into account the fact that he may get factors of production at a lower price than the market transaction which he supersedes…[120]

If we consider an alternative to the firm structure his point is easily understood. Imagine a scenario in which someone building an automobile must repeatedly shop at the local market for material and labor. Initially he buys a frame then returns to purchase successively the wheels, engine, transmission, and so on. Between each acquisition he goes to the labor market and hires temporary labor, first to assemble the wheels on the frame and then later hires more labor to mount the engine, still later more to install doors, windows, etc. This is an extremely inefficient and costly system because of the transaction costs of repetitive visits to the market. Add to this situation consideration of the cost to efficiency incurred because of the need to train or retrain a labor force that turns over with such rapidity we can understand Coase's objection. The rise of the firm is then attributed to this key difference between the transaction costs of obtaining resources in the market and organizing those same resources within the company. If by such an organization the entrepreneur provides his product to the market at a cheaper price than the competition, or the market, and thereby increases the profit the firm generates we can understand the motivation behind the drive to create and operate a firm. Using Coase's definition of the firm we can then say that a firm "…consists of the system of relationships which comes into existence

[120] Coase, "The Nature of the Firm," 392. Here the term 'marketing' takes on a different meaning than in conventional usage, it refers to the costs borne by the company to ascertain the market price and transaction costs of obtaining resources from the market.

when the direction of resources is dependent on an entrepreneur."[121] The directing of resources that Coase refers to is what we call the production process, which we turn to now and examine.

Production Process

Arthur Thompson in *Economics of the Firm: Theory and Practice* defines the production process as "…a series of activities by which resource inputs (raw material labor, capital, land utilization, and managerial talents) are transformed over some period of time into outputs of goods or services."[122] This definition covers a broad range of economic activities that create value irrespective of whether the process is goods-oriented or service-oriented, or whether the enterprise is a not-for-profit or profit making organization.[123] This section details the characteristics, conditions, and assumptions made about the production process. Topics include the nature and characteristics of, and assumptions about, inputs and outputs; the role of technical knowledge; an enumeration of various scenarios coupling inputs with outputs; and lastly, a schema for classifying production. We begin with inputs and outputs.

Inputs and outputs

Inputs and outputs lie at the core of the production process. Inputs are any good or service which make a positive contribution to the production process.[124] For an input to hold relevance or have meaning in an economic analysis of production the input must be scarce and essential to the production process. Inputs are considered essential if the production process

[121] Ibid., 394.
[122] Arthur A. Thompson, *Economics of the Fiirm: Theory and Practice*, Third ed. (Englewood Cliffs, New Jersey: Prentice-Hall, 1981). 154-55.
[123] Ibid.
[124] James M. Henderson and Richard E. Quandt, *Microeconomic Theory: A Mathematical Approach*, Third ed. (New York, New York: McGraw-Hill Book Company, 1980). 64.

requires some positive amount of the corresponding inputs such that all inputs are required in the production process and exist in non-negative quantities. Scarcity exists when inputs are in limited supply such that they have a price greater than zero. Essential but non-scarce inputs, such as sunshine or rainfall, are not included in economic analyses.[125] Finally, scarce but non-essential inputs are likewise eliminated from consideration.

The quantity of inputs utilized in a production process is a function of the underlying technology and the time horizon used in the analysis of that process. Economists consider two time-frames in the analysis of production, the short-run when quantities of some inputs vary according to the technology and level of output while others remain constant or fixed regardless of those same output levels and the long-run when all inputs are considered variable.[126] C. T. K. Ching in *Production Economics* elaborates on the concept of the time-frame used in economic analysis, "A unit of time is said to be short enough so that the entrepreneur cannot alter the quantities of the fixed factors. The unit of time is also assumed to be short enough so that technology does not change the shape of (the) production function. Finally, the time period is assumed to be long enough so that the technical production processes can be completed. Thus, the unit time concept is not specific but general."[127]

Inputs and outputs must be measurable and homogeneous. In empirical studies care must be exercised in determining the unit of measure for inputs and outputs as they impact the ease or practicality of making measurements and calculations.[128] Inputs and outputs are assumed to be

[125] Sune Carlson, *A Study on the Pure Theory of Production* (New York, New York: Augustus M. Kelly, 1965). 12.
[126] Thompson, *Economics of the Fiirm: Theory and Practice*: 177-78.
[127] Ching and Yanagida, *Production Economics: Mathematical Development and Applications*: 89.
[128] Carlson, *A Study on the Pure Theory of Production*: 12.

homogeneous in that no quality difference exists between successive units of inputs, nor does differentiation exist between successive units of output.[129]

People are the essential, scarce and variable resource for architectural firms. Their labor contributes directly to the production of architectural products and services but is also essential in a variety of overhead functions such as marketing. In the latter case the cost of their labor should be added to other overhead costs to become the firm's fixed cost. In the former case the labor constitutes a variable cost applicable to the end product or service and is represented in production functions as the independent input factors. The variable nature of direct labor stems from the ability of the firm to rapidly add to or subject from the total work force as well as allocating people to or removing them from individual projects. The requirement that inputs be homogeneous would at first glance prove problematic for architectural firms but need not be the case. By acknowledging that the skills and years of experience of individual employees vary greatly we have a clue as to how to differentiate them into homogeneous categories. By exercising care and sound judgment firm managers should be able to establish sub-categories for their technical and support staffs that are relatively homogeneous within each category. The issue of homogeneity of output can be resolved in a similar manner. Architectural products and services possessing similar characteristics should be grouped into individual categories and analyzed separately. While the analysis of each output category can be accomplished utilizing the *n*-input single-output model which is our focus here, the basis for eventually considering the *n*-input *m*-output model derives from this homogeneous in output issue.

---

[129] David N. Hyman, *Modern Microeconomics: Analysis and Applications* (St. Louis, Missouri: Times Mirror/Mosby College Publishing, 1986). 5. Beattie allows that in some theoretical and empirical studies this requirement can be relaxed.

Technical Knowledge

Production processes are organized and inputs combined in a manner determined by the entrepreneur. A basic assumption of production theory is that the entrepreneur has discovered, either through trial and error or through research and development, the most efficient arrangements of the factors of production based upon the entrepreneur's technical knowledge. Sune Carlson writes, "If production is defined as a quantitative process of combining certain given productive services, it is the knowledge of these different possible combinations that we call technical knowledge."[130] Later when we state that the production function reflects the underlying technology we explicitly are referring to the design of the production process based upon the entrepreneur's technical knowledge. The design or organization of the production process may vary according to the level of output, i.e. two production lines versus one, but we assume that the underlying technology remains unchanged.[131] Sune Carlson writes "A change in the technical organization of the productive services which accompanies a change in the output must not, of course, be confused with a change in the technical knowledge. A change in the former is a reversible process; a change in the latter is not. … for every service combination, there exists one and only one optimal organization and only one maximum output. A change in technical knowledge, on the other hand, implies that the optimal organization and maximum output from *the same* service combination have changed. "[132] Hal Varian suggests two additional properties of technology in *Intermediate Microeconomics*: *A Modern Approach* that of monotonicity and convexity.[133] In a monotonic technology if at least one input is increased, it should be possible to produce at least the same output as produced originally. In a convex

[130] Carlson, *A Study on the Pure Theory of Production*: 7.
[131] Ibid., 14-15.
[132] Ibid., 15-16.
[133] Hal R. Varian, *Intermediated Microeconomics: A Modern Approach*, Seventh ed. (New York, New York: W. W. Norton & Company, 2006).

technology it may be possible to create the same output with different combinations of inputs but the weighted average of two such sets must be able to produce at least the amount of the original output if not more.[134] For the purpose of this dissertation we will assume that the underlying technology remains unchanging, that it is monotonic and convex. We will return to these last two properties later.

Input-output scenarios

The complexity of the production process dictates the internal organization of the firm. A firm may produce a single product from a single process utilizing a single unique input set; this is the classical *n*-input single-output scenario. Some production processes produce more than one output, termed joint production, from the same input set; examples include meat processing where more than one cut of meat is produced from the same animal. When multiple outputs are obtained from multiple input sets we have the *n*-input, *m*-output scenario. A firm might employ only one single-product single-input set production process to produce a single or joint final output, or many single-product single-input set production processes operating in series or parallel to produce a number of single or joint final products. In complex production operations single or joint production processes may feed their intermediate product (which might also be a final product) to other production processes in a series of processes to complete the final product. These various conditions or scenarios reduce to only four combinations for analysis purposes. First is the single-input single-output process. Second, we have the *n*-input single-output problem. Third, we have the *n*-input joint-output problem. And lastly, we have the more general *n*-input *m*-output problem. In this dissertation we will restrict our investigation to the single product *n*-input case recognizing that multiple parallel processes likely exist within the

[134] Ibid., 326.

firm and that future research may reveal that application of the *n*-input *m*-output model to architectural firms is appropriate.

Thompson Production Classification Scheme

The Thompson scheme for classifying production processes consists of four categories: Unique product production, rigid mass production, flexible mass production, and process production.[135] In unique product production the output is made to a unique set of specifications. Production begins with the receipt of an order and output volume is generally low. The resulting product is relatively unique or special. The process is labor intensive requiring employees that possess high technical skills. By contrast, rigid mass production utilizes and produces highly standardized products in high volume. Employee skill levels vary from high for design and maintenance functions to low or moderate for labor skills used in the actual production line. Flexible mass production utilizes standard products but assembles them in a variety of ways to create diverse products. The skill sets required are similar to those used in rigid mass production. Process production is a continuous activity utilizing a steady flow of resources and converts them into an equally constant flow of outputs. These processes are often highly automated and capital intensive but requiring smaller amounts of labor than the other types of production. The production process employed by architectural firms, and more generally by professional service firms, is the unique product production scheme. Because of the uniqueness of individual projects researchers and analysts applying production theory to architectural production processes must exercise care in grouping similar, yet to a degree divergent, designs or products into narrowly defined product categories. The use of highly

[135] Thompson, *Economics of the Fiirm: Theory and Practice*: 156-58.

skilled employees proficient in various functions within the production process supports our use of multiple categories for labor as opposed to aggregating labor in a single category.

Production Theory

Production economists study the production processes of firms to develop an understanding of the technical and economic characteristics of production systems employed in transforming inputs into goods and services demanded by consumers or other firms.[136] Production theory assumes the optimizing behavior of rational producers as seeking to maximize profit, minimize cost, or achieve optimal allocation of resources under constrained conditions. Thus production theory is about choice among alternatives and how the technical and economic characteristics of the production system influence choices.[137] At the heart of production theory is the production function. The following section defines the term *production function* and provides a detailed discussion of its various characteristics including productivity properties, production curves and surfaces, stages of production, returns to scale, factor substitution, and isoquant analysis among other topics. The section that concludes this chapter is devoted to various optimizing situations and their supporting calculations. No discussion of production theory can precede without first a thorough examination of what is meant by the term *production function.*

Definition of production function

Bruce R. Beattie and C. Robert Taylor in *The Economics of Production* define a production function, "A production function is a quantitative or mathematical description of the

[136] Pappas and Brigham, *Managerial Economics*: 201.
[137] Bruce R. Beattie and C. Robert Taylor, *The Economics of Production* (New York, New York: John Wiley & Sons, 1985). 1.

various technical production possibilities faced by a firm. The production function gives the maximum output(s) in physical terms for each level of the inputs in physical terms."[138] While this is a suitably succinct definition for most purposes it also requires considerable elaboration for consideration in this dissertation.

The first part of the definition '*a quantitative or mathematical description*' requires the production function to take an explicit functional form. However, the ultimate usefulness of the formulation depends upon its adherence to a set of mathematical restrictions. Thus not all functional forms prove useful in production economics. For *n*-inputs the general form of the production function is represented:

$$y = f\left(x\right) = f\left(x_1, x_2, ..., x_i \mid x_{i+1}, ..., x_n\right), \tag{3.1}$$

where y is the output of the function, $x_1, x_2, ..., x_i$ represent the variable inputs and $x_{i+1}, ..., x_n$ represent the fixed inputs. Commonly the expression for the production function omits the fixed inputs and is represented as:

$$y = f\left(x\right) = f\left(x_1, x_2, ..., x_n\right). \tag{3.2}$$

A great many specific forms of the production function exist and these forms are discussed in chapter 4. For immediate illustrative purposes four commonly encountered functional forms are introduced here. The most classical form is the Cobb-Douglas production function,

$$y = A\prod_{i}^{n} x_i^{\alpha_i}, \tag{3.3}$$

or as more commonly represented, [139]

---

[138] Ibid., 3.
[139] A conscious effort is made in this dissertation to utilize a consist methodology in the use of symbols. See Appendix B Mathematical Notations for a complete list of notations used in this dissertation. Many functional forms appear different here than in their original published works. Readers of technical papers are no doubt familiar with the great variety of symbols and notations used by a vast group of writers and are somewhat frustrated by the

$$y = AK^{\alpha}L^{\beta}. \tag{3.4}$$

The other functions are the linear production function,

$$y = \sum_{i}^{n} \beta_i x_i , \tag{3.5}$$

the Leontief production function,

$$y = \min\left[\beta_1 x_1, \beta_2 x_2, \cdots, \beta_n x_n\right], \tag{3.6}$$

and the ACMS constant elasticity production function,

$$y = A\left[\beta_1 x_1^{-\rho} + \beta_2 x_2^{-\rho}\right]^{\frac{-r}{\rho}}. \tag{3.7}$$

These forms are introduced now so that they may later aid in explaining some of the concepts of this chapter or, as in the case of the Leontief production function, demonstrate an exception to the general rule. In addition to taking an explicit form, production functions must adhere to a set of mathematical restrictions which are discussed next.

In *Applied Production Analysis* Robert Chambers gives six assumed properties for production functions.[140]

1. Production functions are quasi-monotonic or strict-monotonic. Under these conditions the addition of any input cannot decrease the level of output. This is equivalent to saying that all marginal products are greater than zero. Later in this chapter when the stages of production are introduced it will be necessary to revisit this property as one characteristic of producing in stage three is the presence of negative marginal products.

---

lack of consistency. Here is a case in point. Normally the Cobb-Douglas function is written, $y = AK^{\alpha}L^{\beta}$ clearly the notation used above is an equivalent form.

[140] Robert G. Chambers, *Applied Production Analysis* (Cambridge, England: Cambridge University Press, 1988). 8-18.

2. Production functions are concave and their input set is convex.[141] Concavity is the necessary condition for the production function to exhibit diminishing marginal productivity. A convex input set implies that if input x_1 and x_2 combine to produce a given output then any weighted average of the two inputs must also produce at least that level of output. Also a convex input set is necessary for the validity of the law of the diminishing marginal rate of technical substitution (MRTS).

3. All inputs are strictly essential. In addition to the earlier stated condition that only those inputs to the production process that are also scarce are of interest in economics, the requirement of strict essentiality requires that the quantity of each input must be greater than zero. Restated, it is not possible to produce a positive level of output without the commitment of each scarce resource. Conversely, a positive output level cannot be obtained without the consumption of a strictly positive amount of each input.

4. The input set is closed and non-empty for $y > 0$. This condition implies that a positive output can always be produced and that discontinuities in the input set are not allowed.

5. $f(x)$ is finite, nonnegative, real valued, and single valued for all nonnegative and finite inputs.

6. $f(x)$ is everywhere continuous and everywhere twice-continuously differentiable.

The second part of the definition of production functions, '*various technical production possibilities*', refers to the underlying technology of the firm. At issue is the difference between technical efficiency and allocative efficiency. Given a range of possible technologies economists assume that entrepreneurs always choose a technology that produces maximum output and

[141] See Appendix C, Concavity and Convexity for a description of the test for concavity or convexity

achieves maximum efficient use of available inputs.[142] However, within the limits imposed by the technology and given the continuous nature of the production function, infinite combinations of inputs can produce a given level of output.[143] The selection of a specific input combination depends upon the relative prices of inputs and outputs.[144] Therefore, the critical consideration is one of allocative efficiency over technical efficiency as the underlying technology is assumed to be the most technically efficient.[145]

The third part of our definition, '*gives the maximum output*', sets a limitation and a restriction. As a limit it establishes the maximum output that a given set of inputs can produce. As a restriction it implies that although a larger input set could produce the same output as a smaller set, the entrepreneur always selects the smaller set as it is the most efficient.

Abstraction is an issue not normally addressed by economists. Production functions are abstractions or models of the real world. Economists often use models or abstractions to simplify their analysis or as aids in explaining theory. The level of abstraction derives directly from the intended use of the model. Production functions used in higher macro level analysis of economies of nations or states often employ aggregated inputs of capital and labor. Similar application of production functions in analysis of the behavior of the firm calls for more disaggregation of capital and labor, or other inputs, into smaller more discrete inputs and that intermediate processes be explicitly represented. Therefore, in addition to the requirements noted above we now must add that the level of abstraction contained in the production function

[142] Ching and Yanagida, *Production Economics: Mathematical Development and Applications*: 88-89.
[143] The Leontief production function is an exception to this rule.
[144] Henderson and Quandt, *Microeconomic Theory: A Mathematical Approach*: 66.
[145] Harvey Leibenstein, "Allocative Efficiency vs. 'X-Efficiency," *The American Economic Review* 56, no. 3 (1966): 392.

accurately represents the economic activity under investigation and that the level of abstraction be appropriate for statistical estimation and practical analysis.[146]

Moving beyond the technical definition of production functions we can now explore various characteristics of production processes represented by or embedded within the production function beginning with the its productivity properties.

Productivity Properties

The theory of marginal diminishing returns can now be told by illustration of the marginal, average and total product concepts and graphs therein derived. Marginal productivity theory states, "…as the amount of a variable input is increased by equal increments and combined with a specified amount of fixed inputs, a point will be reached (sometimes more quickly and sometimes less quickly) where the resulting increases in the quantity of output will get smaller and smaller."[147] This point, corresponding to input level x_1^0 in figure 2 below, is the point of diminishing return. As the variable input is increased from 0 to x_1^0 for each additional unit of variable input output increases by a greater amount than was obtained by adding the previous unit of input. Beyond the point of diminishing returns, but short of diminishing total output, each additional unit of input increases total output, however, the increase from each succeeding unit of input is less than from what was obtained by adding the previous unit of input. At the point of diminishing total output, x_1^1 in figure 2, the increase in total output from the last additional unit of input equals zero. Beyond x_1^1 total output decreases. These relationships are more easily seen in the total product curve shown below in Figure 2.

[146] More about this requirement is contained in the section on estimation.
[147] Thompson, *Economics of the Fiirm: Theory and Practice*: 183.

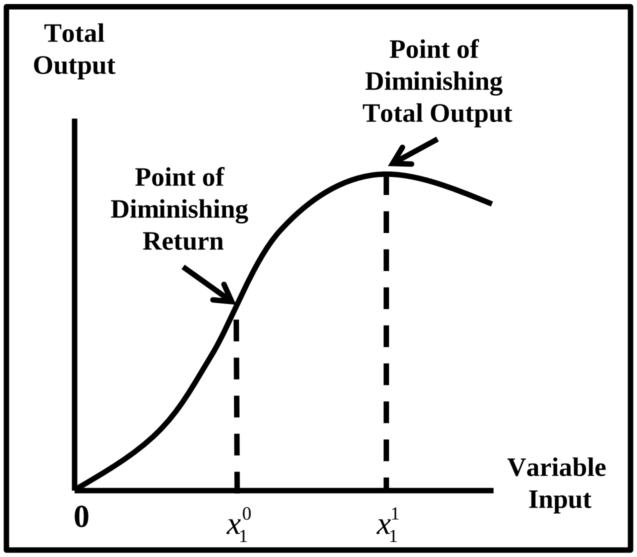

Figure 2, Total Product Curve

Between 0 and x_1^0 the application of additional units of the variable input increases total output and does so at an increasing rate. From x_1^0 to x_1^1 additional units of the variable input continue to increase total output but now does so at a decreasing rate. After the point x_1^1 an addition of a unit of the variable input results in a decline in total output. The Leontief production function, equation(3.6), denies these phenomena by claiming both a constant returns to scale and that the property of diminishing returns does not apply.

Marginal Product

Marginal product is the amount of change in total output for each additional unit of variable input. For the single variable input case of the type $y = f(x) = f(x_1)$ the first derivative of the function, denoted as f', gives the marginal product. For the *n*-input case, $y = f(x) = f(x_1, x_2, \ldots, x_n)$, the marginal product for each variable input is given as f_i i=1 to n, where f_i is the first partial derivative of $f(x)$ with respect to the *i*th input. In figure 3 below, the marginal product is positive and increasing in the range 0 to x_1^0, between x_1^0 and x_1^1 it remains positive but is declining, and after x_1^1 becomes negative.

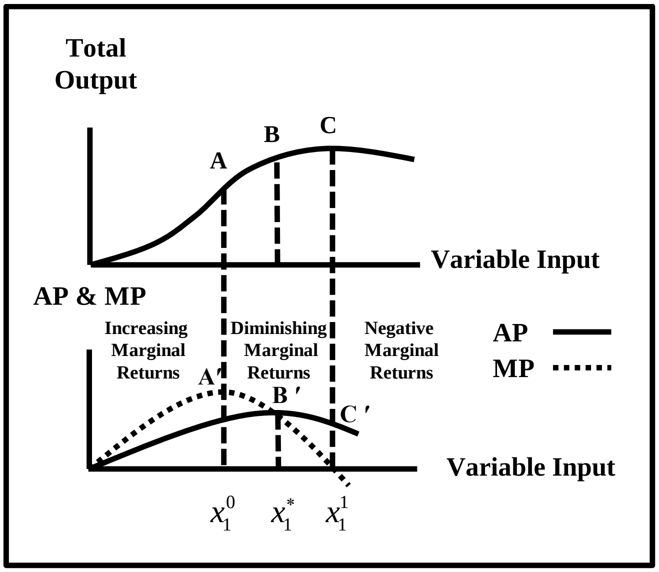

Figure 3, Average, Marginal, and Total Product Curves

Average Product

The average product is given by $AP = \frac{f(x)}{x}$ for the single input case and by $AP = \frac{f(x)}{x_i}$ for i=1 to *n* for the *n*-input case. A graphic representation of the relationship between marginal product (MP) and average product (AP) is found in Figure 3, above. We see that until the level of the variable input reaches x_1^0 that marginal product and average product increase. The point of diminishing marginal returns exists at the points A and A′. From variable input level x_1^0 to x_1^* the marginal product begins to decrease while average product continues increasing. At the points B and B′ average product peaks then begins to decline, this is the point of diminishing average product. Between variable input levels x_1^* and x_1^1 both average product and marginal product decline but remain positive. At the point C and C′ total product peaks and thereafter declines; this is the point of diminishing total product. After this point average product and marginal product decline with marginal product becoming negative.

Production Surface

By extending the single variable example used in graphing total production, as depicted in figures 2 and 3, to a two variable input example we can graph the production surface as shown in figure 4 below, with the variable input x_1 and x_2 plotted on the horizontal plane and output on the vertical axis.

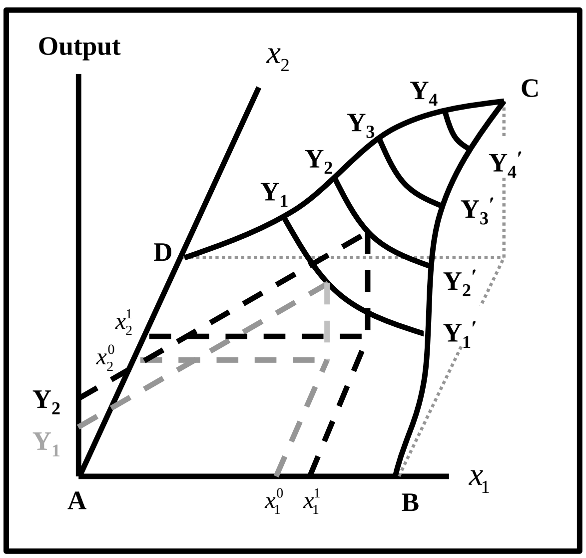

Figure 4, Production Surface

The curve DC represents the production curve of x_1 when x_2 is held fixed and conversely the curve BC represents the production curve of x_2 when x_1 is held constant. By allowing both to vary a production surface ABCD is revealed. The slope of curves BC and DC are the marginal products of their respective inputs. In plotting specific values of x_1^0 / x_2^0 or x_1^1 / x_2^1 the output levels Y_1 / $Y_1{}'$ and Y_2 /$Y_2{}'$ respectively are revealed. In the *n*-input case the production surface cannot be directly constructed; however, by constructing a series of graphs of related pairs of inputs their relationship may be discerned and an overall characterization of the production function be arrived at and prove useful.[148]

Isoquants

The output levels (Y_1, Y_2, etc.) depicted in figure 4 are termed isoquants. An isoquant depicts the locus of all combinations of any two inputs, i.e. x_1 and x_2, that produce a specified output level; it is analogous to the consumer's indifference curve.[149] Flattening the graph in figure 4 produces the family of isoquants curves shown in Figure 5. Each isoquant represents a different level of output with the smallest output closest to the origin while each isoquant moving away from the origin represents higher levels of output such that $y_1 < y_2 < y_3 < y_4$.

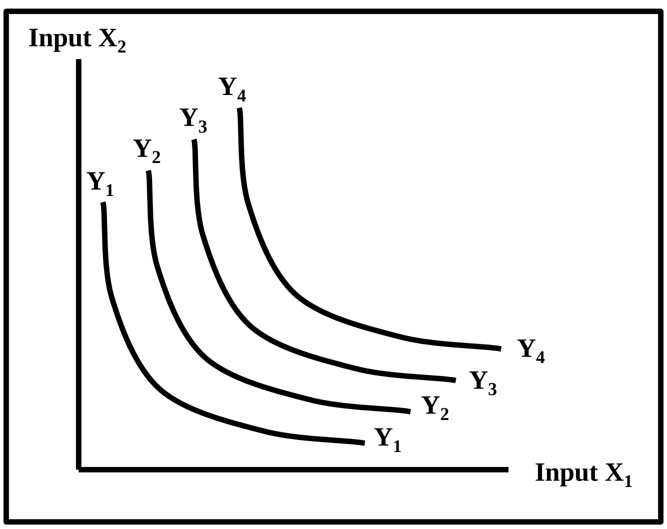

Figure 5, Isoquants

---

[148] Carlson, *A Study on the Pure Theory of Production*: 18-19.
[149] Henderson and Quandt, *Microeconomic Theory: A Mathematical Approach*: 70.

In *Intermediate Microeconomic Theory* David Kamerschen provides 5 characteristics that isoquants must exhibit. [150]

1. Isoquants are continuous and everywhere dense. This condition is necessary to provide for substitution between inputs.

2. Isoquants are negatively sloping. The segment of the isoquant curve of economic interest must exhibit a positive marginal product. While mathematically an isoquant curve potentially displays both positive and negative marginal products we are only interested in the viable economic region where the marginal product is positive.

3. Isoquants are convex to the origin. Similar to the requirement that the isoquant be negatively sloped, convexity is necessary to maintain the validity of the law of diminishing marginal rate of technical substitution in the viable economic region.

4. Isoquants are non-intersecting.

5. Isoquants do not cut the axis of an essential input.

The formula for the isoquant derives from the production function. In the two input case the production function is given as,

$$y = f\left(x_1, x_2\right). \tag{3.8}$$

To find the equation for the isoquant we take the inverse of the function with respect to x_2, preferring that x_1 remains on the horizontal axis, and obtain,

$$x_2 = f^{-}\left(x_1, y\right). \tag{3.9}$$

Using the Cobb-Douglas production of the form,

$$y = Ax_1^{\alpha_1} x_2^{\alpha_2}, \tag{3.10}$$

[150] David R. Kamerschen and Lloyd M. Valentine, *Intermediate Microeconomic Theory* (Cincinnati, Ohio: South-Western Publishing, 1977). 192-94.

we find the inverse to be,

$$x_2 = A^{1/\alpha_2} x_1^{-\alpha_1/\alpha_2} y^{1/\alpha_2}. \tag{3.11}$$

Thus given the parameters $A, \ \alpha_1, \ \alpha_2$ we can, for a specified output *y*, vary x_1 and find the appropriate value for x_2. To find additional isoquants we need only further vary the specified output level. The usefulness of isoquants is not limited to a simple depiction of output levels as the next section illustrates.

Isoquant Analysis

Isoquant analysis supports economic analysis of the firm as it reveals a number of useful characteristics of a production function and the underlying technology. Through analysis of the slope, shape, and spacing of isoquants we are able to discern the relationship between the various inputs and their impact on output levels. This analysis reveals the marginal rate of technical substitution; provides a means to confirm the convexity of isoquants; exposes the degree of substitution inherent in the underlying technology; and graphically represents the type of returns to scale.

Marginal Rate of Technical Substitution[151]

The marginal rate of technical substitution (MRTS) describes the manner in which two inputs substitute for each other and is the slope of the isoquant. Simply stated the MRTS shows the amount of one factor that must be sacrificed per incremental increase in the other factor. Thus the MRTS is the ratio of the change in both input factors;

---

[151] Beattie and Taylor, *The Economics of Production*: 23-24.

$$\text{MRTS} = \frac{dx_2}{dx_1}. \tag{3.12}$$

Using the total differential of the production function we can derive a formula for the value of the MRTS. The total differential of the production function in the two input case is,

$$dy = f_1 dx_1 + f_2 dx_2. \tag{3.13}$$

As the change in output along an isoquant is zero the total differential becomes,

$$0 = f_1 dx_1 + f_2 dx_2. \tag{3.14}$$

By simplification of terms we obtain,

$$\frac{dx_2}{dx_1} = -\frac{f_1}{f_2}. \tag{3.15}$$

Thus the MRTS is the ratio of the first partial derivatives of x_1 and x_2. In the *n*-input case it is necessary to examine the inputs in a pair wise fashion to determine all relevant MRTS's.

Economists generally drop the negative sign and merely refer to the MRTS as the ratio $\frac{f_1}{f_2}$.

Convexity in Isoquants[152]

Confirmation of the convexity of the isoquant is arrived at by taking the second derivative of the MRTS, where in all cases f_i is the first partial derivative of the production function with respect to input *i* and f_{ij} is the cross partial derivative with respect to inputs *i* and *j*, and f_i^2 is the square of the first partial derivative of input *i*.

The second derivative of the total differential is,

[152] Ibid., 24-25.

$$d\left(\frac{dx_2}{dx_1}\right) = \frac{\partial\left(-\frac{f_1}{f_2}\right)}{\partial x_1}dx_1 + \frac{\partial\left(-\frac{f_1}{f_2}\right)}{\partial x_2}dx_2. \tag{3.16}$$

Dividing both sides by dx_1 we obtain,

$$\frac{d\left(\frac{dx_2}{dx_1}\right)}{dx_1} = \frac{d^2x_2}{dx_1^2} = \frac{\partial\left(-\frac{f_1}{f_2}\right)}{\partial x_1} + \frac{\partial\left(-\frac{f_1}{f_2}\right)}{\partial x_2}\frac{dx_2}{dx_1}. \tag{3.17}$$

Wishing to hold the output *y* constant we need to substitute $-\frac{f_1}{f_2}$ for $\frac{dx_2}{dx_1}$, and by using the quotient rule for partial derivatives we obtain,

$$\frac{d^2x_2}{dx_1^2} = -\frac{f_2 f_{11} - f_1 f_{21}}{f_2^2} - \frac{f_2 f_{12} - f_1 f_{22}}{f_2^2}\left(-\frac{f_1}{f_2}\right). \tag{3.18}$$

By multiplying the first term by $\frac{f_2}{f_2}$ and completing the multiplication of the second term we obtain,

$$\frac{d^2x_2}{dx_1^2} = \frac{-f_2^2 f_{11} + f_1 f_2 f_{21} + f_1 f_2 f_{12} - f_1^2 f_{22}}{f_2^3}, \tag{3.19}$$

and by collecting terms arrive finally at,

$$\frac{d^2x_2}{dx_1^2} = \frac{-f_2^2 f_{11} + 2 f_1 f_2 f_{21} - f_1^2 f_{22}}{f_2^3}. \tag{3.20}$$

The isoquants of the production function are convex provided the sign of this second derivative is positive.

Some production functions exhibit the unique characteristic of possessing positive and negative marginal products thus creating isoquants with areas of convexity and concavity as shown in Figure 6. Only in area I of figure 6 are the marginal products all positive and the isoquant convex. In area II the isoquants are positively sloped and convex with respect to the vertical axis. The isoquants in area III are concave to the origin. Lastly in area IV the isoquants are positively sloped and convex with the horizontal axis. Due to this variability in isoquant behavior it is necessary to restrict the function to area I. This is consistent with the view that a rational entrepreneur will only operate where marginal products are positive and isoquants are convex to the origin. The areas in figure 6 lying outside of area I represent irrational choices for any output produced by input combinations lying on the isoquant in those areas could easily be produced by fewer inputs if produced by combinations laying on the isoquant in area I.[153]

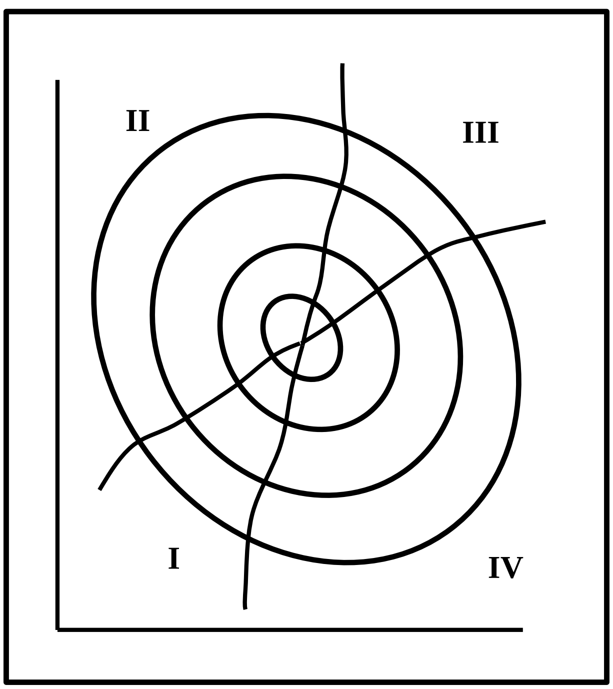

Figure 6, Convexity and Concavity of Isoquants

[153] Ibid., 26-29.

Factor Substitution

Factor substitution and its converse factor separability are essential characteristics of production functions. Factor substitution describes the degree or manner in which one input or factor can be exchanged for another factor. Elasticity of factor substitution is the customary measure of factor substitution; however, isoquants can be used to graphically depict the degree of substitution.

Elasticity of Factor Substitution

Elasticity measures the proportional change in one quantity resulting from a change in another quantity. In factor substitution the elasticity of factor substitution (σ) measures the rate at which substitution takes place. σ measures the change in the input ratio divided by the proportionate change in the marginal technical rate of substitution and is given by,[154]

$$\sigma = \frac{f_1 f_2 \left(f_1 x_1 + f_2 x_2 \right)}{x_1 x_2 \left(2 f_{12} f_1 f_2 - f_1^2 f_{22} - f_2^2 f_{11} \right)}. \tag{3.21}$$

σ can take a value between zero and infinity, $0 \leq \sigma \leq \infty$. When the elasticity of substitution equals zero the inputs are not substitutes for another and must be used in strict fixed proportions. The Leontief production function, equation(3.6), is an example of a production function that utilizes inputs in a fixed proportion or recipe fashion. When the elasticity of substitution equals infinity the inputs of the production function are perfect substitutes for one another. The linear production function of equation (3.5) is an example of a production function with perfect substitution. Most production functions exhibit imperfect substitution, a situation where inputs compete for their use in the production process. Under imperfect substitution the elasticity will

[154] See Appendix D for the derivation of the elasticity of substitution.

vary between zero and infinity. The Cobb-Douglas production function, equation(3.3), is an example of a production function exhibiting imperfect substitution.[155]

Factor Substitution Graphically

The shape of the isoquant curve can also reveal the nature of factor substitution. Isoquants taking the shape of those depicted in graph A of figure 7 below are representative of production functions producing standard isoquants associated with imperfect competition between inputs. The Cobb-Douglas production function, equation(3.3), is an example one such production function. Graph B in figure 7 depicts the straight line isoquants produced by production functions exhibiting perfect substitutability. The linear production function, equation(3.5), exhibits perfect substitution. Graph C of figure 7 illustrates the bent isoquants associated with complementary inputs or the no substitution scenario. The Leontief production function, equation(3.6), is an example of this type of production function.[156]

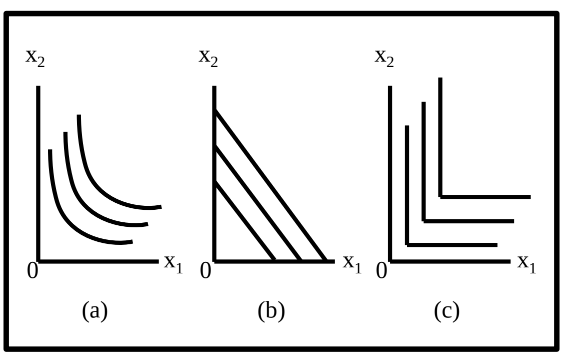

Figure 7, Shapes of Isoquants

[155] Beattie and Taylor, *The Economics of Production*: 29-31.
[156] Ibid.

Stages of Production

Economists find it convenient to break the total product curve into three stages of production based upon the changing nature of both the average and marginal products. As shown in figure 8 below stage one is characterized by an increasing average product curve with an initially increasing marginal product curve that begins to decline beyond the point of diminishing marginal return. Stage one ends at the point where marginal product equals average product. Stage two is characterized by declining average and marginal product curves with both exhibiting positive values. Stage two ends at the point of diminishing total product; while it exhibits a positive average product, marginal product has declined to zero. In stage three marginal, average, and total product curves decline with marginal product becoming negative. Absent knowledge of the prices of inputs or outputs the rational entrepreneur would prefer to operate where the greatest technological efficiency exists which is at the threshold of stage one and two if the variable input is the scarcest. Should the fixed input be the scarcest then the threshold moves to the boundary of stage two and three. Given knowledge of prices entrepreneurs will operate somewhere within stage two.[157]

[157] Ching and Yanagida, *Production Economics: Mathematical Development and Applications*: 103-05.

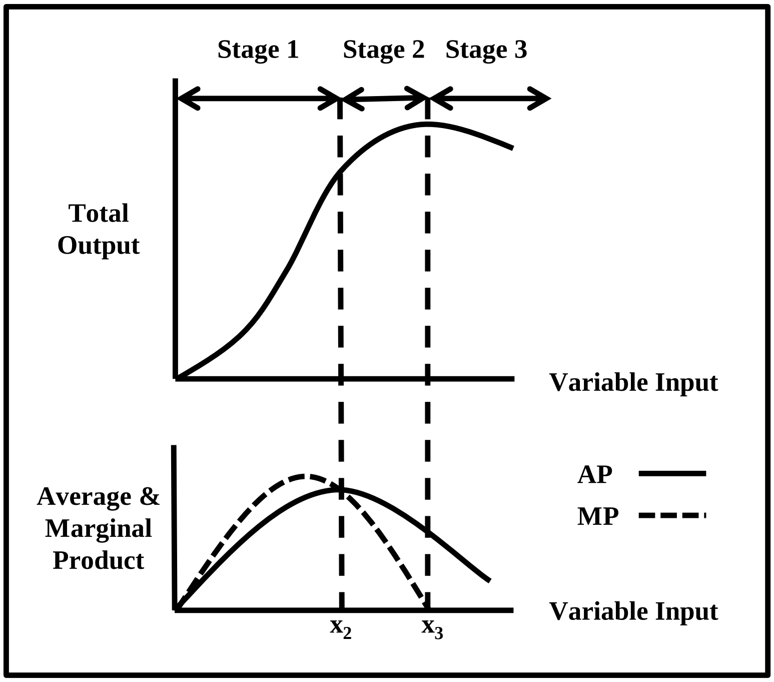

Figure 8, Stages of Production

Isoclines and Ridgelines

Isoclines are special features of the isoquant map that locate and display points of equal slope of a family of isoquants. Two special cases of isoclines are ridgelines and the expansion path. For ordinary isoclines the slope of interest is the marginal technical rate of substitution. The formula for the isocline is derived from the MRTS which for the two input case is $\frac{f_1}{f_2}$. By setting this expression equal to a constant value of the MRTS we obtained the formula $\frac{f_1}{f_2} = K$. Given that f_1 and f_2 are both functions of x_1 and x_2 we can express the formula in terms of x_2 and by varying x_1 obtain the graph of the isocline. By selecting various values for K a family of isoclines is obtained. Likewise for the *n*-input case we need only select pairs of x_i's and repeat the process to obtain a family of isoquants and associated isoclines. Figure 9 shows a family of isoclines, ridgelines, and an expansion path. All the lines A-F in figure 9 are isoclines; we can see that isoclines may be straight (C), curved (A, B, D, E) or wavy (f). Ridgelines are a special type of isocline as the associated slope is either zero or undefined (i.e. the MRTS is zero or undefined). Ridgelines, lines A and E in figure 9, effectively graph the boundaries of the area of stage II production where the marginal products are all positive. A second special case of the isocline is the expansion path. Entrepreneurs are interested in knowing the least-cost resource combinations for several potential output levels. Given knowledge of the ratio of input prices and setting that value as the slope K in our isocline formula the resultant isocline is called the expansion path. Thus the expansion path, line F in figure 9, locates the various input combinations for different levels of output provided we know the relevant input prices. The expansion path may be wavy as in line F of figure 9, straight as in line C, or curved as in lines B and D.[158]

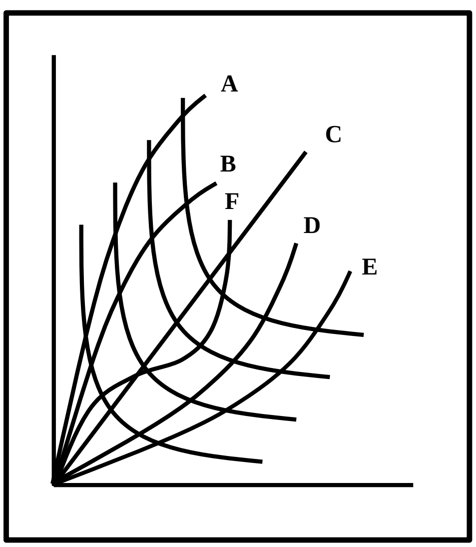

Figure 9, Isoclines, Ridgelines, and Expansion Path

[158] Beattie and Taylor, *The Economics of Production*: 31-32.

Elasticity of Production

As noted earlier elasticity measures change. In output levels the elasticity of production measures the proportional change in output resulting from a proportional change in a given factor, holding the level of other inputs constant. The formula for the elasticity of production is given by the ratio of the marginal product of the *i*th input to its average product,

$$E_i = \frac{f_i}{\left(\frac{f(x)}{x_i}\right)} = \frac{f_i x_i}{y}. \tag{3.22}$$

Stages of Production Redux

We can now return to the stages of production and provide an alternative explanation of the boundaries of the stages of production. Figure 10 below is the same as figure 8 except that now we use values of the elasticity of production for x_1 noting that input levels for x_2 are held constant. Stage one production exists in the region of the graph between o and x_1^0 where $E_1>1$ or the condition when marginal product is greater than average product. The boundary between stage one and two production exists where $E_1=1$, or the condition that marginal product equals average product. Stage two exists in the region of the graph where E_1 lies between zero and one or when the condition when average product is greater than marginal product and marginal product remains positive. The boundary between stage two and three exists where $E_1=0$ or the condition that marginal product equals zero. Stage three exists in the region where $E_1<0$ or the condition that marginal product is negative.[159]

[159] Ibid., 36-38.

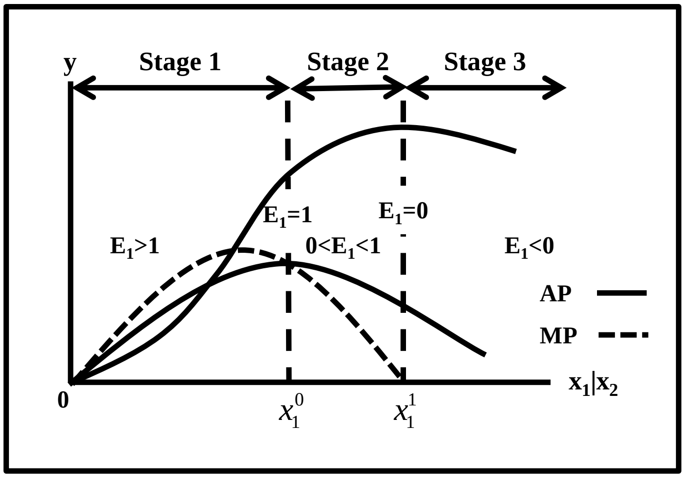

Figure 10, Stages of Production

Coefficient of Production

By contrast to the elasticity of production where the level of one input is allowed to vary while holding the other inputs constant, the function coefficient measures the change in output given an equal increase in all inputs. Thus the function coefficient reveals the returns to scale of the production process. For the two-input case the function coefficient is given by the formula,

$$\varepsilon = \frac{f_1 x_1}{y} + \frac{f_2 x_2}{y}. \tag{3.23}$$

From equation (3.22) we see that for the two-input case equation (3.23) can be replaced by,

$$\varepsilon = E_1 + E_2. \tag{3.24}$$

Thus the function coefficient is equal to the sum of the individual elasticities of production. Expanding the formula to the *n*-input case then the function coefficient is given by,

$$\varepsilon = \sum_{1}^{n} E_i. \tag{3.25}$$

When ε is less than one the production process exhibits declining returns to scale, when ε equals one it exhibits constant returns to scale, and when greater than one the process exhibits increasing returns to scale. Some production functions demonstrate only one type of return to scale and are referred to as fixed return to scale production functions, while others have characteristic of all three types of return to scale and are referred to as variable return to scale production functions.[160]

Isoquants and Return to Scale

Isoquants can graphically demonstrate returns to scale. In figure 11 below the spacing of the isoquants reflect the three types of returns to scale. In graph (a) we have increasing returns to scale. Because a unit increase in all inputs generates more than a unit increase in output the isoquants spread out as output grows. Graph (b) shows constant returns to scale. In constant returns to scale a unit increase in all inputs results in a unit increase in output thus the isoquants are equally spaced. Graph (c) show the isoquants associated with decreasing returns to scale. Each unit increase in inputs results in less than a unit increase in output thus the isoquants grow increasing bunched as output grows.

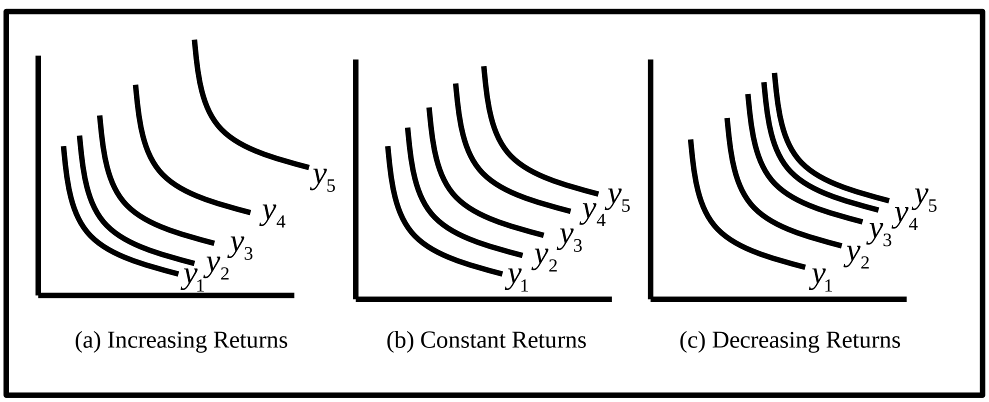

Figure 11, Isoquants and Returns to Scale

[160] Ibid., 40-48.

Factor Interdependence

Factor interdependence expresses a technical relationship between two factors. Two factors are independent of one another if an increase in the quantity used of one factor does not change the marginal product of the other factor. Three technical interdependence states exist; complementary, independent, and competitive. The test for interdependence utilizes the cross partial derivative of the production function with respect to the two inputs in question, or f_{ij}. When f_{ij} is greater than one, the inputs are technically complementary. In the complementary case the marginal product of one input is enhanced as the quantity of the other input is increased in the production process. Inputs are independent if f_{ij} equals zero. In the case of independence the marginal products of both inputs are not affected by changes in the quantity used of the other. In the third case inputs are technically competitive when f_{ij} is less than zero. Under competitive conditions the marginal product of one input is reduced when the other input level is increased.[161] Figure 12 below graphically shows these three relationships.

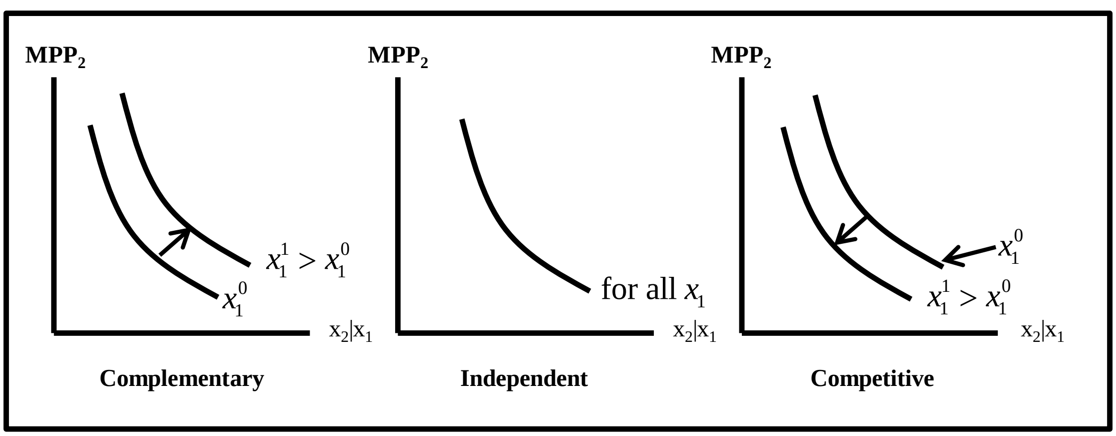

Figure 12, Factor Interdependence

[161] Ibid., 32-36.

The importance of factor interdependence is manifest in applied studies. In model specification selection of functional form becomes critical. The form selected should match the factor interdependence inherent in the production process being model rather than imposing a different type of factor interdependence.[162]

Factor Separability

Factor separability represents a third type of technical relationship between factor inputs, with elasticity of substitution being the first and factor interdependence the second type. The question at issue here is whether an intermediate process and resultant product can be identified and separated from the main process. If an intermediate process can be identified then that process can be represented by its own production function and the inputs of the intermediate process can be aggregated into a single input for the overarching or primal production function. Thus the question to aggregate inputs or disaggregate inputs in the primal production function is raised. Aggregate inputs are often used in the analysis of industrial sectors and national economies.[163] A group of inputs (one or more inputs) are said to be separable from the other inputs if the marginal rate of technical substitution between the group and the remaining inputs are independent.[164] The test for separability utilizes the partial and cross partial differential of the production function. A pair of inputs *ij* is separable from input *k* if and only if the following condition holds,

$$\frac{f_{ij}}{f_i} = \frac{f_{jk}}{f_j}. \tag{3.26}$$

---

[162] Ibid.
[163] Charles Blackorby, Daniel Primont, and R. Robert Russell, *Duality, Separability, and Functional Structure: Theory and Economic Applications* (New York, New York: Elsevier North-Holland, 1978). 5.
[164] Ibid., 1.

In cases where equation (3.26) holds then researchers can specify a functional form for the production function that incorporates this condition.

Production Theory in Use

In concluding this section on production theory we are left with the question of how to put this information to use. The answer is three-fold. First of all this brief statement of production theory conceptualizes production in a manner not existing outside of the field of economics. This understanding permits detailed mathematical analysis of production in architectural firms not previous attempted. The dynamics and intricacies of production revealed here inexorably alter future discourses about issues arising about production processes within the firm. Second, this understanding of production theory permits the more thorough discussion of production functions found in the next chapter and furthers our understanding that it's the very complexity of production that spawns the proliferation of production functions seen in the 21st century. Thirdly, production theory provides managers the tools to answer key questions in their quest to optimize their production operations as we shall see in the following section.

Optimization

The preceding section presented the fundamental elements of production theory focusing on the two-input and *n*-input single-output production models thus serving as a prelude to an exploration of the practical application of that theory. That entrepreneur's exhibit optimizing behavior as an everyday aspect of their management of the firm is perhaps the most fundamental assumption that economists hold true. Such optimizing behavior attempts to answer the questions of 'how do I minimize the costs of a given level of output?' and 'how do I maximize output for a given level of costs?' Thus the two most common optimizing questions involve minimizing cost for a given level of production (constrained cost minimization) and maximizing production for a given budget constraint (constrained output maximization). Given knowledge of input and output prices, that is the price of production factors and the market price of goods and services produced, an entrepreneur armed with production theory can answer these important questions.

In introducing the concepts of optimizing Charles Beightler in *Foundations of Optimization* comments that the goal is that of finding the best solution.[165] When the analysis is one of mathematics, particularly the attempt to find the maximum or minimum value of a function of many variables constrained by an objective function, the solution derived is highly precise, often to several decimal places. Such precision is false and misleading and thus a note of caution should be sounded. While the general form of a production function suggests its own high level precision, the values of all parameters in a specified form are derived from statistical estimation methods. All parameters (constants, coefficients and exponents) represent point estimates of their true value. While these estimates are often expressed to several decimal places each have attached to them their own standard error reflecting the likelihood that the real value exists with a given number of standard deviations from the predicted value. From these standard errors confidence intervals may be constructed to better give the researcher a sense of the true value and probability of an event occurring outside the confidence interval. Therefore a completed statistical analysis of the form of a production function must be viewed with some circumspection. The errors inherent in the production function carry over into optimization calculations and taint their conclusions. The entrepreneur must therefore consider that a solution

[165] Charles S. Beightler, Don T. Phillips, and Douglass J. Wilde, *Foundation of Optimization*, Second ed. (Englewood Cliffs, New Jersey: Prentice-Hall, 1979). 1.

calling for 1000 units of input x_i or an output of 1000 units or that a given input bundle has a total cost of $1,000, has a natural variance from that predicted value and that variance must be taken into account when implementing any solution. Thus the best solution sought by Beightler is at best a little fuzzy; caution is called for.

Caution must be exercised in employing decision-making that relies upon optimizing techniques, particularly mathematical optimization. Beightler offers these words of guidance, "Deciding how to design, build, regulate, or operate a physical or economics system ideally involves three steps: First, one should know accurately and quantitatively, how the system variables interact. Second, one needs a single measure of system effectiveness expressible in terms of the system variable. Finally, one should choose those values of the system variables that yield optimum effectiveness. Thus optimization and choice are closely related."[166] Ergo improving our fuzzy solution requires a little hard work. Let us briefly explore each of these aspects of optimization before proceeding to the technical aspects of this section.

In Beightler's scheme the first requirement – knowledge of the system – is of paramount importance.[167] In the present context two requirements arise, that of knowledge of production theory including economic optimization and knowledge of the specific production system of the architectural firm. The preceding section on production theory provides us the requisite knowledge of the production theory while the balance of this section details optimization techniques. What we must supply next is our specific knowledge of how productive factors interrelate in a given production process. This knowledge manifests itself in model specification and data acquisition; subjects we will deal with later.

[166] Ibid., 2.
[167] Ibid., 3.

Identifying system effectiveness measures is Beightler's second step. He further suggests that such measures may vary from trivially simple to practically impossible.[168] For the architectural firm the immediate measures of effectiveness are those naturally proceeding from the analysis of cost minimization and production maximization. Cost minimization identifies the optimal expenditure conditions (total cost plus ratio of input factors) for a given output. From this benchmark the firm can judge the efficiency in design or other services by comparison of the actual cost with that of the optimal cost and associated variances. One application of the result of such analysis directly informs decisions regarding sizing and composition of the project team necessary in achieving maximum economic efficiency. By setting the output equal to that of projects the firm is bidding upon an estimate of labor cost broken down into the various labor categories is possible and may prove useful as part of the bidding process. Also the future size and composition of the work force may be determined through that analysis by setting the output equal to an annual plan of work. Production optimization produces another benchmark for the firm. Through comparison with actual production (e.g. annual production) with the optimal level the firm may directly assess its level of productivity. In a similar fashion the size and composition of a proposed or future work force may be tested to determine if it can meet anticipated future output levels. While the principal focus remains understanding the cost of design, or other architectural services, clearly much insight and information of value to the firm is obtainable through these optimization techniques.

Choice is the third of Beightler's three steps in optimization. Rational decision making on the part of entrepreneurs remains a leading assumption of economists. Beightler asserts that such rational behavior is consistent with optimizing and in fact should be based upon solutions suggested by optimizing techniques as they narrow the number of possible solutions down to the

[168] Ibid.

best one.[169] Additionally he notes, “Moreover, it often yields information about the sensitivity of optimum conditions to fluctuations and uncertainties in the original system description.”[170] In the two optimization techniques described in the balance of this section we find the key to answers of many of the decisions managers face.

Optimizing Behavior

The search for optimality is an effort to find a singular point where the cost function, represented by isocost curves, and production functions, represented by isoquant curves, yield the best solution for the optimization question at hand. The optimal point exists where the isocost line and the isoquant curve are tangent thus establishing a point of minimum cost for a fixed output or maximum output for a given cost. The isoquant curves introduced earlier represent various input combinations yielding a fixed output as depicted in figure 5. Isocost lines depict various combinations of inputs having the same total cost as shown in figure 13. The formula for total cost of a two input system is,

$$c_0 = w_1 x_1 + w_2 x_2 + FC, \tag{3.27}$$

where c_0 is the total cost, w is the wage rate, and FC represents fixed costs.[171] For the *n*-input case the cost formula is generalized to,

$$c_0 = \left(\sum_{i=1}^{n} w_i x_i \right) + FC. \tag{3.28}$$

By solving for x_1 in equation (3.27) we obtain the following formula,

$$x_1 = \frac{c_0 - FC}{w_1} - \frac{w_2}{w_1} x_2. \tag{3.29}$$

---

[169] Ibid.
[170] Ibid.
[171] Henderson and Quandt, *Microeconomic Theory: A Mathematical Approach*: 74.

The slope of the isocost line is $-\frac{w_2}{w_1}$ which is the negative of the ratio of the input prices, while

$\frac{c_0 - FC}{w_1}$ is the intercept on the x_1 axis.[172] In figure 12 we see three isocost lines, c_1, c_2 and c_3, each line consisting of various input bundle combinations that have the same total cost such that the total cost for $c_1 < c_2 < c_3$.

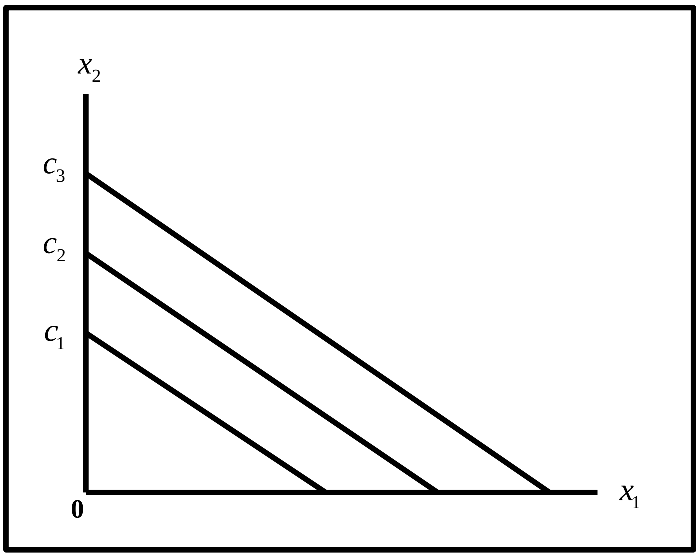

Figure 13, Isocost Lines

The optimal point exists where an isocost line and isoquant curve are tangent. In figure 14 the points A, B, and C show the point of tangency between isocost lines c_1, c_2, and c_3 and isoquants y_1, y_2 and y_3 respectively. These optimal points can be arrived at mathematically through use of the Lagrangian formulation which we explore next.

[172] Ibid.

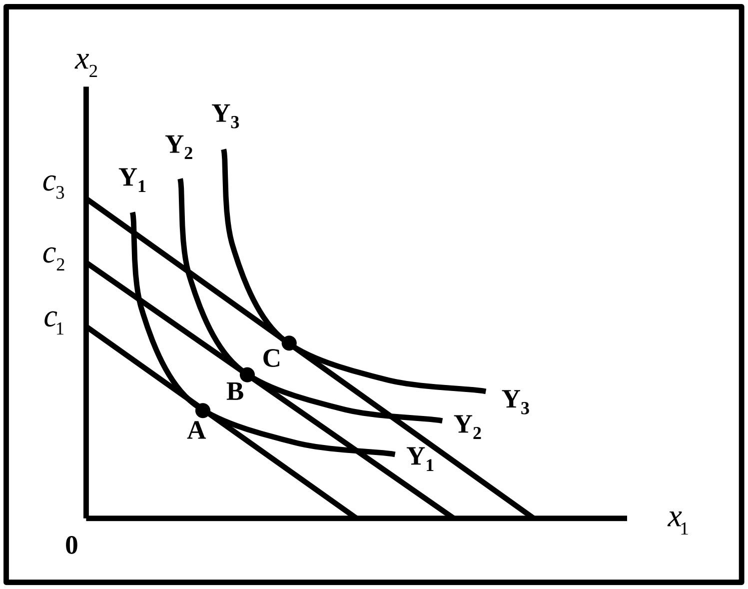

Figure 14, Isocost Lines and Isoquant Curves

Constrained Cost Minimization[173]

One question faced by entrepreneurs is how to minimize the cost of producing at a given production level. This optimization problem is termed constrained cost minimization. For the two input case we seek to minimize the cost function in equation (3.27) subject to the production function. For the two-input case we begin by forming the Lagrangian function,

$$L = w_1 x_1 + w_2 x_2 + FC + \lambda\left(y_0 - f\left(x_1, x_2\right)\right) \tag{3.30}$$

Next we must determine the first order conditions of the minimization problem. Recall that in univariate calculus the minimum is found by setting the first derivative equal to zero and solving for x. Here the process is similar only we have three variables x_1, x_2, and λ, and we are dealing with partial derivatives of the Lagrangian function, $f\left(L\right)$. Thus to establish the first order conditions begin by setting the partial derivatives of L with respect to x_1, x_2, and λ equal to zero and obtain,

[173] Ibid., 76-78. This entire section is based on the explanation contained in these pages.

$$\frac{\partial L}{\partial x_1} = w_1 - \lambda f_1 = 0 \tag{3.31}$$

$$\frac{\partial L}{\partial x_2} = w_2 - \lambda f_2 = 0 \tag{3.32}$$

$$\frac{\partial L}{\partial \lambda} = y_0 - f\left(x_1, x_2\right) = 0\,. \tag{3.33}$$

Next solve for the variables in equations (3.31) and (3.32) to obtain,

$$w_1 = \lambda f_1 \tag{3.34}$$

$$w_2 = \lambda f_2\,. \tag{3.35}$$

The λ term in equation (3.30) becomes 1 with differentiation thus yielding,

$$y_0 = f\left(x_1, x_2\right). \tag{3.36}$$

By dividing equation (3.34) by (3.35) we find,

$$\frac{w_1}{w_2} = \frac{f_1}{f_2}\,. \tag{3.37}$$

Recalling from earlier that $\frac{f_1}{f_2}$ equals the marginal rate of technical substitution we now conclude that the first order conditions require that the ratio of the prices of the factors of production equal the MRTS.

The second order conditions require that the determinant of the bordered Hessian matrix to be positive. For the two input case the following must be true to ensure a valid optimal condition.

$$\begin{bmatrix} f_{11} & f_{12} & -w_1 \\ f_{12} & f_{22} & -w_2 \\ -w_1 & -w_2 & 0 \end{bmatrix} > 0\,. \tag{3.38}$$

The procedure for the *n*-input case is very similar. The Lagrangian formulation is,

$$L = \left(\sum_{i=1}^{n} w_i x_i \right) + FC + \lambda \left(y_0 - f\left(x_1, x_2, \ldots, x_n\right) \right). \tag{3.39}$$

Solving the Lagrangian function for the first order conditions we obtain *n* functions for the variables $x_1 \ldots x_n$ similar to equations (3.31) and (3.32), plus a function for λ similar to equation (3.33). The second order conditions are also similar only with an expanded bordered Hessian matrix of the following type,

$$\begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1n} & -w_1 \\ f_{12} & f_{22} & \cdots & f_{2n} & -w_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ f_{1n} & f_{2n} & \cdots & f_{nn} & -w_n \\ -w_1 & -w_2 & \cdots & -w_n & 0 \end{bmatrix} > 0. \tag{3.40}$$

Appendix E presents an example of the two input case demonstrating the techniques to reach a final solution of the constrained cost minimization problem.

Constrained Output Maximization

The second most frequently asked question by managers is how to produce the maximum output given a predetermined budget. This optimization problem is termed constrained output maximization. For the two input case we seek to maximize output given the constraint of a predetermined budget limit. The procedure is similar to the constrained cost minimization problem. We start with the Lagrangian formula,

$$L = f\left(x_1, x_2\right) + \lambda \left(C_o - w_1 x_1 - w_2 x_2 - FC \right). \tag{3.41}$$

The first order conditions are obtained by differentiating the Lagrangian function with respect to each variable and setting them equal to zero,

$$\frac{\partial L}{\partial x_1} = f_1 - \lambda w_1 = 0 \tag{3.42}$$

$$\frac{\partial L}{\partial x_2} = f_2 - \lambda w_2 = 0 \tag{3.43}$$

$$\frac{\partial L}{\partial \lambda} = C_0 - w_1 x_1 - w_2 x_2 - FC = 0\,. \tag{3.44}$$

Next solve for the variables in equations (3.43) and (3.44) to obtain,

$$\lambda = \frac{f_1}{w_1} \tag{3.45}$$

$$\lambda = \frac{f_2}{w_2}\,. \tag{3.46}$$

In differentiating the function L with respect to λ the λ term becomes 1 once again thus yielding,

$$C_0 = w_1 x_1 + w_2 x_2 + FC\,, \tag{3.47}$$

which is the definition of the cost curve. As the left hand term in both equations (3.45) and (3.46) is λ we can rewrite those equations as:

$$\frac{f_1}{w_1} = \frac{f_2}{w_2} \tag{3.48}$$

and by simple algebra obtain,

$$\frac{f_1}{f_2} = \frac{w_1}{w_2} \tag{3.49}$$

thus demonstrating again that the first order conditions require the MRTS to equal the ratio of the input prices.

The second order conditions are identical in both optimization problems with the determinant of the bordered Hessian matrix required to be greater than zero. (See equation(3.38)) An example of the two input case is demonstrated in Appendix E.

The *n*-input case for the constrained output maximization proceeds similar to the constrained cost minimization procedure noted above.

The foregoing chapter contains the exposition of production theory for the *n*-input single-output model, a detailed examination of the production process and characteristics of production functions then concluded with a detailed explanation of two optimizing techniques. With the foundation formed by these discussions we are now prepared to examine the vast array of production functions developed in the 20th century to address the various production conditions discussed in this chapter.

CHAPTER 4

PRODUCTION FUNCTIONS

Economics is extremely useful as a
form of employment for economists.
John Kenneth Galbraith[174]

Introduction

This chapter surveys production functions developed in the 20th Century focusing upon those production functions that take explicit form and are of the *n*-input single-output variety. This survey enumerates these production functions, provides a brief history of their origin and salient characteristics. The enumeration aspect of this survey provides a rough chronological order of development of production functions beginning with the Cobb-Douglas production function. The enumeration deviates from a strict chronology where clarity suggests grouping developments in one place. Otherwise the intent is to detail the development of production functions from the early part of the 20th century up to the present. While exhaustive this enumeration suffers the fate of most enumeration efforts. As Melvyn Fuss and Daniel McFadden state the issue, "The diversity and extent of the subject of applied production theory makes a comprehensive survey impossible."[175] What follows in this chapter can, none-the-less, provide an acceptable start in the search for appropriate forms of production functions suitable for empirical research into the production processes employed by architectural firms.

As noted in chapter two, in the early 18th Century works of economists one can find conceptualizations of production functions including the isolated appearance of actual mathematical formulations of production functions. Unfortunately, much of that work was not widely published nor circulated. That which was published, often in very small quantities, was

[174] Thinkexist.com, "Thinkexist.com," www.thinkexist.com/quotes/john_kenneth_galbraith.
[175] Fuss and McFadden, *Production Economics: A Dual Approach to Theory and Applications, Vol I and II*: 219.

not well known to a great many economists. Differences in language and the distances between great centers of learning often hindered the sharing of knowledge. Regrettably the early work of economists often failed to gain widespread acknowledgement or review. A critiquing of that early work and the eventual acceptance of concepts articulated therein awaited their rediscovery many years after their original publication. By the early part of the 20th century much of the contemporary work on production economics became centered in the United States, in no small part due to the immigration of notable economists to American universities. Improvements in communication technology meant that more of their scholarship could and would be shared and critiqued widely. These developments set the stage for the wide spread exchange of ideas and acknowledgment of the contributions of a plethora of U.S. and international economists. The work of such luminaries as Sune Carlson, Charles Cobb and Paul Douglas, Wassily Leontief, and Paul Samuelson, to name just a few, headlined the early effort of economists in America. The salient event marking the beginning of this new contemporary period in economics was the 1928 publication of *The Theory of Production* by Charles Cobb and Paul Douglas in which the Cobb-Douglas production function was first revealed.[176] In the main section of this chapter we take up the enumeration of production functions beginning there with Professors Cobb and Douglas and their famous function.

A simple enumeration of production functions does not fully forward our main objective in this dissertation, namely the identification production functions suitable for empirical research into the operations of architectural firms. Accomplishing that goal requires an investigation into the origins, purpose, and characterization of explicit production functions of the *n*-input single-output variety. So in addition to the enumeration of each production function a short description of its salient features are provided. These features include an account of the manner in which the

[176] Cobb and Douglas, "A Theory of Production."

production function handles the issues of factor separability and interaction, the measurement of returns to scale, the applications for which it was developed, and its generality and flexibility. These factors are discussed further in the first section of this chapter.

Readers unfamiliar with production economics will, no doubt, be astonish by the virtual cornucopia of functions surveyed here. One may be tempted to ask why so many? The balance of this chapter sheds considerable light upon the answer to this question without attempting to fully resolve it. Economists were often critiquing or responding to ideas of earlier economists proposing their own solutions to old problems or presenting new issues as they identified them. Economists advanced differing treatments for issues ranging from input separability to returns to scale; the adaptation of production functions to a broader array of production circumstances; and factor substitutability to name only a few prominent issues. Some economists focused upon the development of production functions for high level macroeconomic analysis, others focused upon functions for microeconomic applications, while still others sought to develop production functions with broader applications. As the development of production functions continued the issues of generalizing the form of the function, and the development of flexible forms arose. Consequently we have today an extensive array of production functions designed to confront an equally broad array of conditions and circumstances. Alongside the enumeration of production functions the intent in the next section is a brief comment upon the unique characteristics of each. The purpose of this characterization is to aid in the eventual task of winnowing a large list of possibilities down to a manageable few worthy of use in empirical economic analysis of architectural firms.

General Characteristics of Production Functions

Introduction

This section discusses the four principle means of classifying production functions used in the main section of this chapter. These classifications are centered upon the issues of homogeneity; elasticity and general substitutability/interaction of inputs; the economic application for which it was developed; and the degree to which the form generalizes early forms, extends them to multiple inputs and provides for flexibility in empirical analysis.

Homogeneity

The first characteristic of production functions that commands our attention is that of homogeneity. The homogeneity of a production function immediately informs us of the nature of the returns to scale of that function. In the discussion of returns to scale in chapter three it is noted that returns to scale of a production function may be increasing, decreasing or constant. Further, fixed return to scale production functions are those functions whose return to scale does not change over the relevant data range although that return to scale itself may be increasing, decreasing, or constant. Variable returns to scale production functions refer to those functions whose return to scale do vary over the relevant data range. In many production functions the return to scale is assumed to be fixed over the relevant economic range of data. For instance the Cobb-Douglas production function may prove to be increasing, decreasing, or constant, but regardless of what the return to scale is determined empirically to be, it remains fixed over the relevant data range. In other words once the returns to scale is determined empirically it remains invariable or fixed. Variable returns to scale implies that over some range of inputs the production function may exhibit increasing returns to scale, such as when small amounts of inputs are employed, then changes to a constant return to scale (or very nearly constant return to scale) as more inputs are applied, before ultimately exhibiting decreasing returns to scale.[177] The homogeneity of the production function reveals the nature of the returns to scale with homogeneous production functions exhibiting fixed returns to scale (increasing, decreasing or constant), while non-homogeneous production functions exhibit variable returns to scale (e.g. changing from increasing to constant to decreasing). For homogeneous functions the extent to which the function is increasing (or decreasing) may easily be determined and for non-homogeneous functions the returns to scale must be determined for specified level of inputs.

A production function is defined as a homogeneous function if, when the amount of each input used is multiplied by a scalar, λ, the output y changes by, λ^k.[178] Thus the function $y = f\left(x_1, x_2\right)$ is homogeneous in degree k if $y\lambda^k = f\left(\lambda x_1, \lambda x_2\right)$.[179] If $k > 1$ then the function is increasing in returns to scale and as the value of k increases the degree of expansion of y increases proportionally. If $k = 2$ for example, then a doubling of inputs results in a four-fold increase in output. If $k = 1$ then the function exhibits a constant return to scale. When $0 < k < 1$ then the function exhibits decreasing returns to scale and as k decreases output decreases proportionally. A non-homogenous function is one that fails the test for homogeneity. The returns to scale of such functions are typically variable.

Factor Elasticity, Substitution, Separability, and Interaction

The second characteristic of production functions to command our attention is how does a specific function account for elasticity of factor substitution, factor independence, factor

---

[177] Henderson and Quandt, *Microeconomic Theory: A Mathematical Approach*: 105-06.
[178] Ching and Yanagida, *Production Economics: Mathematical Development and Applications*: 101.
[179] Henderson and Quandt, *Microeconomic Theory: A Mathematical Approach*: 106.

separability, and factor interaction. As discussed in chapter three the measure of the elasticity of substitution of a production function describes the manner or degree in which inputs may be substituted one for another. This ranges from perfect substitution when $\sigma = \infty$ where all inputs are fully substitutable for each other, to imperfect substitution when $0 < \sigma < 1$ where inputs are substitutable based upon the rate of technical substitution, to the condition of non-substitutability when $\sigma = 0$. Similar to the discussion above regarding the returns to scale, the measure of factor substitution may be fixed in some production functions while held variable in others. In modeling a production process the assumptions regarding the manner of substitution often prove critical.

Economic Application

The third characteristic to command our attention concerns how production functions are employed in economic analysis. The two areas we are concerned with here are macroeconomic and microeconomic applications. Many production functions developed for use in studying national economies, industrial sectors, or aggregated production are not suitable for microeconomic applications. In many instances they allow only the aggregation of capital and labor and cannot be extended to handle multiple inputs required in microeconomic analysis. Similarly some microeconomic analysis utilizes functions that account for sophisticated interactions among biological agents or general agricultural circumstances and thus may not be appropriate for certain manufacturing processes. Of interest in this study are those production functions that can easily be extended to handle multiple inputs.

Generalized and Flexible Forms

The last characteristic to command our attention is the trend toward development of production functions that encompass a number of special treatments under one formulation. The first such type of production function is the generalized functional form. The generalized form extends a previously developed production function to include more inputs or to account for more variability than the original form. Flexible forms are intended as 'one size fits all' forms. These are often nested functions in that as certain parameters tend toward zero, one, or infinity. Or are assigned a predetermined value, the flexible form resolves itself into one of the more specific functions encompassed within the flexible form. Such a function might for example contain parameters that permit the flexible form to have contained within it other functions such as the Cobb-Douglas and Leontief functions simultaneously. Only upon statistical estimation does one know which of the included functions emerges. Flexible forms come with a cost in the estimating process. These functions often cannot be estimated with ordinary least square methods thus requiring sophisticated techniques or complex software programs. The added parameters serve not only to complicate the estimation process but increase the number of values that must be estimated along with their standard errors thus complicating interpretation of the results. Finally, an increase in the number of parameters that must be estimated increases the minimum number of observations required for statistical validity. As we shall see in the next chapter the simpler or more parsimonious function is greatly preferred when estimating the values of the production function.

Productions Functions

Cobb-Douglas Production Function

In the spring of 1927 Paul Douglas[180] began a study "to determine what relationships existed between the three factors of labor, capital, and product" in U.S. manufacturing.[181] He turned to his colleague at Amherst College, Charles W. Cobb of the mathematics faculty, and asked "if he could devise a mathematical function which could be used to measure the comparative effect of each of the two factors upon the total product."[182] The result of this famous collaboration is the Cobb-Douglas production function[183] of the form, [184]

$$y = Ax_1^{\alpha_1}x_2^{\alpha_2} \mid \alpha_1 + \alpha_2 = 1. \tag{4.1}$$

This original formulation requires the sum of the exponents α_1 and α_2 to equal one thus making the function homogeneous in degree one and thus a fixed constant return to scale function. Additionally the function exhibits imperfect substitution. The optimal equilibrium point is found where the MRTS equals the ratio of the respective wage rates. Thus the manner in which substitution of one input for another depends upon both the MRTS and prevailing wages. Originally devised as a macroeconomic analysis tool it has gained most of its noteworthiness as a microeconomic tool. This is the first production function that introductory students of microeconomics gain exposure. In microeconomic applications its more general form is often encountered,

---

[180] Murray Brown, ed. *The Theory and Empirical Analysis of Production*, vol. 31, Studies in Income and Wealth (New York, New York: National Bureau of Economic Research and Columbia University Press, 1967), 15.
[181] Cobb and Douglas, "A Theory of Production," 139.
[182] Brown, *The Theory and Empirical Analysis of Production*, 16.
[183] Cobb and Douglas, "A Theory of Production," 152.
[184] A conscious effort is made in this dissertation to utilize a consist methodology in the use of symbols. See Appendix B Mathematical Notations for a complete list of notations used in this dissertation. Many functional forms appear different here than in their original published works. Readers of technical papers are no doubt familiar with the great variety of symbols and notations used by a vast group of writers and are somewhat frustrated by the lack of consistency. Here is a case in point. In the original the formulation appeared as $P' = bL^kC^{1-k}$.

$$y = A\prod_{1}^{n} x_i^{\alpha_i} \ . \tag{4.2}$$

In 1937 David Durand proposed a modification to the original formulation contending that the exponents need not sum to one.[185] This removes the qualification in the original formulation for the exponents to equal one thus allowing for values greater or less than one. Under these conditions the Cobb-Douglas remains homogeneous but no longer only in degree one. Now the degree of homogeneity can be greater than one, indicating increasing returns to scale, or less than one, indicating decreasing returns to scale, or remain one as in the original formulation. Regardless of which type of returns to scale the formula takes upon estimation it is fixed over the relevant data range.

Leontief Production Function

In 1942 Wassily Leontief derived a fundamentally different approach to the production function resulting from his work on input-output models of general equilibrium. That work led him to believe that inputs were related in fixed proportions or recipe style arrangements. Using the example of baking a cake this can be easily illustrated. Baking a cake requires a minimum amount of each ingredient in a strict ratio to the other ingredients. If this balance is upset the result is an inedible concoction. A baker with half the required amount of flour may bake a smaller cake using the other ingredients in the proper proportions but cannot bake a regular size cake regardless of how much sugar or eggs she has on hand. In this same way the Leontief production function assumes fixed proportions in the combination of inputs. For a production process yielding one unit of output from a combination of one unit of input ‘A’ and 2 units of

---

[185] David Durand, "Some Thoughts on Marginal Productivity, with Special Reference to Professor Douglas' Analysis," *Journal of Political Economy* 45, no. 6 (1937): 755.

input ‘B’ an abundance of either, but not both, inputs still results in one unit of output. Should that process have 6 units each of both ‘A’ and ‘B’ it can make only 3 units of output with 3 units of input ‘A’ left over. The Leontief formulation takes the following form,[186]

$$y = \min(\beta_1 x_1, \beta_2 x_2) . \tag{4.3}$$

The Leontief production function is homogeneous in degree one thus is a fixed constant return to scale function. Substitution of inputs is not permitted as $\sigma = 0$. In the cake baking example sugar cannot be substituted for flour. While different types of flour can be utilized they presumably produce different types of cakes, not the original cake, and may be present in the mix in different proportions than the original recipe. In application the Leontief was originally devised for macroeconomic purposes but has been widely employed in microeconomic applications as well. Like the Cobb-Douglas it is easily extended to accommodate multiple inputs as follows,

$$y = \min\left(\beta_1 x_1, \beta_2 x_2, \cdots, \beta_n x_n\right) . \tag{4.4}$$

Later we will encounter attempts to generalize the original formulation and encompasses it in broader based flexible forms.

Halter, Carter, and Hocking (HCH) Transcendental Production Function

In *A Note on the Transcendental Production Function* in 1957 A. N. Halter, H. O. Carter, and J. G. Hocking[187] introduced the transcendental[188] production function. In the conclusion of

---

[186] Wassily W. Leontief, *The Structure of American Economy, 1919-1929: An Empirical Application of Equilibrium* (Cambridge, Massachusetts: Harvard University Press, 1941). 34-41.
[187] A. N. Halter, H. O. Carter, and J. G. Hocking, "A Note on the Transcendental Production Function," *Journal of Farm Economics* 39, no. 4 (1957).
[188] Paul Erdos and Underwood Dudley, "Some Remarks and Problems in Number Theory Related to the Work of Euler," *Mathematics Magazine* 56, no. 5 (1983). The term transcendental derives from the use of the transcendental number e, Euler is often credited with defining a transcendental number as one that is not the root of an algebraic expression.

their article they indicated the motivation behind its development "The function described and illustrated by the foregoing appears to be useful in describing data that show the three traditional phases of the marginal product curve …"[189] The transcendental production function takes the form,

$$y = Ax_1^{\alpha_1} e^{\gamma_1 x_1} x_2^{\alpha_2} e^{\gamma_2 x_2} \cdots x_n^{\alpha_n} e^{\gamma_n x_n} . \tag{4.5}$$

In its general form that translates to,

$$y = A \prod_1^n x_i^{\alpha_i} e^{\sum_1^n \gamma_i x_i} . \tag{4.6}$$

The HCH transcendental production functions is a non-homogeneous function which, depending upon the values of the data and estimated parameters, may exhibit any or all types of returns to scale. Due to its multiplicative nature (similar to the Cobb-Douglas production function) its elasticity of substitution ranges $0 \leq \sigma \leq 1$ thus exhibiting imperfect and variable elasticities of substitution between inputs. Originally developed in response to agricultural settings the function is equally applicable to microeconomic applications as it can easily handle multiple inputs. While the formulation appears messy estimating the parameters is rather straight forward given a log transformation and yields to ordinary least square estimation techniques. However, estimating this function for small data sets may prove problematic as the number of parameters to be estimated is $2p+1$, where p represents the number of inputs. This is the first flexible production function we encounter for when all the $\gamma_{i's} = 0$ the function takes the form of the Cobb-Douglas production function, equation(4.2).

[189] Halter, Carter, and Hocking, "A Note on the Transcendental Production Function," 974.

Heady and Dillon (HD) Production Functions

The array of available production functions expanded significantly in 1961 with the publication of *Agricultural Production Functions*, by Earl Heady and John Dillon based upon their work on agricultural systems at Iowa State University.[190] Incorporating the theory of production economics into their studies of agricultural production they developed an extensive collection of production functions. The common characteristic of these formulations is their connection to agricultural growth studies. Unlike industrial production functions noted above, the basis for substitution in these new production functions has to make allowance for non-linear interaction effects. In the Cobb-Douglas substitution, an imperfect substitution scheme, some portion of the work producing the output is assumed by the input replacing or substituting for the initial input without an increase in product y. A more complex interaction is assumed in the production functions to follow. Take the simple example of adding nutrients 'A' and 'B' to the soil. If the effects of these two nutrients are largely independent aside from their relative effectiveness in stimulating growth we are free to substitute, as in the Cobb-Douglas function, upon the combined effects of their individual influences and their relative costs. If, however, a more complex interaction exists between the inputs a more complex formulation is needed. Take the example again of nutrients 'A' and 'B'. How do we account for the case when even a small amount of either nutrient magnifies the impact of the other nutrient? In the functions that follow these interactions are specifically allowed for in the formulation of the function. It is not clear that such relationships occur or do not occur in labor intensive applications such as the architectural firm. One case where such interaction might occur is in the relationship between mentor and intern where a small amount of supervision reaps large dividends.

[190] Earl O. Heady and John L. Dillon, *Agricultural Production Functions* (Ames, Iowa: Iowa State University Press, 1961).

HD-"Quadratic" The form of the "quadratic" function is:

$$y = A + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2 + \beta_{21} x_2 x_1 \tag{4.7}$$

In its extended form it becomes,

$$y = A + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} x_i x_j \ . \tag{4.8}$$

HD-"Cubic" The form of the "cubic" function is,

$$y = A + \beta_1 x_1 + \beta_2 x_2 + \delta_1 x_1^2 + \delta_2 x_2^2 - \gamma_1 x_1^3 - \gamma_2 x_2^3 \, . \tag{4.9}$$

In its extended form it becomes,

$$y = a + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \delta_i x_i^2 + \sum_{i=1}^{n} \gamma_i x_i^3 \ . \tag{4.10}$$

HD-"Square Root" The form of the "square root" function is,

$$y = A + \beta_1 x_1 + \beta_2 x_2 + \delta_1 x_1^{.5} - \delta_2 x_2^{.5} + \gamma_{12} x_1^{.5} x_2^{.5} \, . \tag{4.11}$$

In its extended form it becomes,

$$y = A + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \delta_i x_i^{.5} + .5 \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} x_i^{.5} x_j^{.5} \ . \tag{4.12}$$

HD-"One and a Half Power" The form of the "One and a Half Power" function is,

$$y = A + \beta_1 x_1 + \beta_2 x_2 + \delta_1 x_1^{1.5} + \delta_2 x_2^{1.5} + \gamma_{12} x_1 x_2 \, . \tag{4.13}$$

In its extended form it becomes,

$$y = A + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \delta_i x_i^{1.5} + .5 \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij} x_i x_j \ . \tag{4.14}$$

HD-"Inverse" The form of the "Inverse" function is,

$$y = \frac{1}{A} + \beta_1^{-1} X_1 + \beta_2^{-1} x_2 + \beta_{12}^{-1} x_1 x_2 \, . \tag{4.15}$$

In its extended form it becomes,

$$y = \frac{1}{A} + \sum_{i=1}^{n} \beta_i^{-1} x_i + .5 \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij}^{-1} x_i x_j \tag{4.16}$$

All of the Heady-Dillon production functions are non-homogeneous and exhibit variable returns to scale. Each employs a slightly different substitution scheme and recognizes a variety of interaction parameters. Originally developed for agricultural and biological applications, their application in microeconomic analysis of the firm has not been established. In modeling any of these functions an empirical researcher must be careful in understanding the dynamics of the workplace in order not to be misled by statistical results. Additionally, the number of parameters that must be estimated expands rapidly as the number of inputs increases. While statistical estimation may prove worthwhile with a small number of independent variables as that list grows the estimation process quickly becomes unmanageable.

Newman-Read Production Function

Also appearing in 1961 is a generalized Cobb-Douglas production proposed by P. K. Newman and R. C. Read in *Production Functions with Restricted Input Shares*.[191] Aggregation studies of national production have long held that the relative share of national product between capital and labor is virtually constant.[192] The Newman-Read production function relaxes the requirement that relative input shares be held constant. The two-input form of this function is,

$$y = A x_1^{\alpha_1} x_2^{\alpha_2} e^{\alpha_3 \ln x_1 \ln x_2} . \tag{4.17}$$

The three-input formulation is rather messy and cumbersome to estimate, while the four-input is virtually impractical for modeling and estimation purposes. The three-input form is

---

[191] P. K. Newman and R. C. Read, "Production Functions with Restricted Input Shares," *International Economic Review* 2, no. 1 (1961).
[192] C. E. Ferguson and Ralph W. Pfouts, "Aggregate Production functions and Relative Factor Shares," *International Economic Review* 3, no. 3 (1962): 328.

$$y = Ax_1^{\alpha_1} x_2^{\alpha_2} x_3^{\alpha_3} e^{\left(\alpha_{12} \ln x_1 \ln x_2 + \alpha_{13} \ln x_1 \ln x_3 + \alpha_{23} \ln x_2 \ln x_3 + \alpha_{123} \ln x_1 \ln x_2 \ln x_3\right)}. \qquad (4.18)$$

The Newman-Read production function is a non-homogeneous function whose return to scale factor is dependent upon the values of the respective inputs and parameters. Developed as a generalized form of the Cobb-Douglas production function its primary application is in the realm of macroeconomic analysis of national aggregation studies. It is a flexible form to the extent that when $\alpha_3 = 0$ in Equation (4.17) or all $\alpha_{ij's}$, and $\alpha_{ijk's} = 0$ in Equation (4.18) the result is the standard Cobb-Douglas function.

Arrow, Chenery, Minhas, Solow (ACMS) Production Function

The Newman-Read production function addressed the issue of input share constancy. The article titled *Capital-Labor Substitution and Economic Efficiency*[193] authored by K. J. Arrow, H. B Chenery, B. S. Minhas and R. M. Solow extends that discussion by noting the unsatisfactory consequence of the two then prevailing assumptions regarding substitution, namely the constant input coefficients of the Leontief function and the unitary elasticity of substitution found in the original Cobb-Douglas function.[194] Citing empirical evidence they established a case for values of σ other than one and that this value varied among a range of products. In response they developed the constant elasticity of substitution (CES) production function, or ACMS, whose form is,[195]

$$y = A\left(\beta x_1^{-\alpha} + (1-\beta) x_2^{-\alpha}\right)^{-\frac{1}{\alpha}}. \qquad (4.19)$$

[193] Kenneth J. Arrow et al., "Capital-Labor Substitution and Economic Efficiency," *The Review of Economics and Statistics* 43, no. 3 (1961).
[194] Ibid., 225.
[195] The derivation of this production function can be found in Appendix F. The general process demonstrated for the first time in deriving this production function has been employed since then by many economists.

The parameter A is an efficiency factor in that as A increases output increases proportionally; the β parameter is the distribution or factor share parameter; and the α parameter is the elasticity of substitution factor, its value determines the returns to scale factor of the function. The permissible values of α are $-1 \le \alpha \le \infty$ as these values produce isoquants of the correct curvature. As a homogeneous function the elasticity factor determines whether the function's fixed return to scale represents increasing ($-1 \le \alpha < 0$), constant ($\alpha = 0$) or decreasing ($0 < \alpha < \infty$) returns to scale. Developed utilizing aggregate capital and labor data the immediate application is one of macroeconomic analysis. Shortly we will encounter adaptations of the basic CES function that permits microeconomic analysis. The ACMS CES model is a flexible form in that as $\alpha \to \infty$ the function reduces to the Leontief model, at $\alpha = -1$ it becomes a linear model, and as $\alpha \to 0$ it becomes the CD production function.[196]

Brown and De Cani (BD) Production Function

Working independently but in parallel with the efforts of Arrow-Chenery-Minhas-Solow, M. Brown and J. S. De Cani developed a CES production function nearly identical to the ACMS CES model but providing for variable returns to scale.[197] The form of the BD production function is as follows,

$$y = A\left(\beta x_1^{-\alpha} + (1-\beta)x_2^{-\alpha}\right)^{-\frac{v}{\alpha}}. \tag{4.20}$$

In this form the function is homogeneous in degree v exhibiting increasing returns to scale when $v > 1$, constant returns when $v = 1$, and decreasing returns when $v < 1$. Unlike the ACMS model

---

[196] Arrow et al., "Capital-Labor Substitution and Economic Efficiency," 230-31.

[197] Murray Brown and John S. De Cani, "Technological Change and the Distribution of Income," *International Economic Review* 4, no. 3 (1963).

the returns to scale of the BD production function varies across the range of data. Like the ACMS model its origins lie in macroeconomic applications of aggregation studies.

Uzawa-McFadden (UM) CES Production Function

In the ACMS and BD CES models, noted above, the elasticity of substitution σ is a function of only two input factors. Due to the symmetry of partial derivatives ($f_1 f_2 = f_2 f_1$), the σ of a two input system is symmetrical ($\sigma_{ij} = \sigma_{ji}$). In an *n*-input system the constancy of the elasticity of substitution between any pairs of inputs and the remaining inputs is questionable. That is, does $\sigma_{12} = \sigma_{13} = \sigma_{23}$ in a three-input example? Working independently and upon slightly different lines of analysis Hirofumi Uzawa[198], in 1962, and Daniel McFadden[199], in 1963, demonstrated the constancy of partial elasticities and successfully extended the ACMS model from two inputs to an *n*-input model. The form of the Uzawa-McFadden CES production function is,

$$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha}\right)^{-\frac{v}{\alpha}}. \tag{4.21}$$

In this form the function is non-homogeneous with the parameter v describing the returns to scale in the normal fashion and α the elasticity of substitution which is constant. With this extension to *n*-inputs the CES function can now be successfully utilized in *n*-input microeconomic analysis.

[198] Hirofumi Uzawa, "Production Functions with Constant Elasticities of Substitution," *The Review of Economic Studies* 29, no. 4 (1962).
[199] Daniel McFadden, "Constant Elasticity of Substitution Production Functions," *The Review of Economic Studies* 30, no. 2 (1963).

Ferguson-Pfouts Production Function

In 1962 C. E. Ferguson and Ralph W. Pfouts extended the discussion regarding the constancy of relative shares, or constant elasticity of substitution, with a publication entitled *Aggregate Production Functions and Relative Factor Shares*[200] wherein they proposed two variations on the CD function to account for changes in relative input shares. These functions are,

$$y = x_1^{\alpha_1} x_2^{1-\alpha_1} e^{\alpha_2 x_1} \tag{4.22}$$

and

$$y = x_1^{\alpha_1 - \left(\frac{\alpha_3}{x_2}\right)} x_2^{\alpha_2}. \tag{4.23}$$

Both of these formulations are non-homogeneous functions possessing the twin characteristics of variable elasticities of substitution and variable returns to scale.[201] Developed for aggregation of national levels of capital and labor they are strongly macroeconomic in applications and to date have not been extended to allow for multiple inputs and therefore are of limited use is microeconomic analysis.

Hildebrand-Liu-Bruno (HLB) Production Function

George Hildebrand and Dazhong Liu published *Manufacturing Production Functions in the United States* in 1965 based upon empirical studies they conducted circa 1962.[202] This work led them to a regression equation pertaining to their research on manufacturing in the United

---

[200] Ferguson and Pfouts, "Aggregate Production functions and Relative Factor Shares."
[201] Aly A. Helmy, "A Family of Generalized Transcendental Production Functions" (Dissertation, University of Nore Dame, 1981), 55.
[202] George H. Hildebrand and Dazhong Liu, *Manufacturing Production Functions in the United States*, vol. 15, Cornell Studies in Industrial and Labor Relations (Ithaca, New York: New York State School of Industrial and Labor Relations, Cornell University, 1965).

States. In an unpublished paper dated 1962 M. Bruno used the regression formula and converted in to the following production function,[203]

$$y = A\left(\beta_1 x_1^{(1-\alpha)\gamma} x_2^{\alpha\gamma} + (1-\beta)\, x_2^{\gamma}\right)^{\frac{1}{\gamma}} \tag{4.24}$$

This formulation is a non-homogeneous function exhibiting variable rates of substitution and returns to scale. Derived from economic studies of nationally aggregated data and not being suitable for extension to more than two inputs its application is limited in microeconomic analysis and is thus principally macroeconomic in nature.

Mukerji Production Function

In a review of the work accomplished by Uzawa and McFadden, V. Mukerji noted that the form of equation (4.21) results in constant and identical partial elasticities of substitution.[204] She viewed this as a defect in the formulation and proposed a correction with the following,

$$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha_i}\right)^{-\frac{1}{\alpha_0}} \tag{4.25}$$

This is a non-homogeneous function except under the condition when $\sum_{i=1}^{n} \alpha_i = 1$. Under this condition and when α tends to zero and $\sum_{i=1}^{n} B_i = 0$, the function reduces to the generalized Cobb-Douglas form. To evaluate the return to scale factor sum the $\alpha_{i's}$ and compare that sum with the value of α_0, when they are equal then constant returns to scale exist, if the sum is greater in

[203]Michael Bruno, *A Note of the Implications of an Empirical Relationship Between Output per unit of Labor, The Wage Rate, and the Capital Labor Ratio* (Stanford, California: Stanford University, 1962).
[204] V Mukerji, "A Generalized S.M.A.C. function with Constant Ratios of Elasticity of Substitution," *The Review of Economic Studies* 30, no. 3 (1963).

represents increasing returns, and if small then decreasing returns exist. This formulation is appropriate for microeconomic studies when the partial elasticities are assumed to be unequal.

Ferguson Transcendental Production Function

In 1965 C. Ferguson proposed the following transcendental function to remedy a defect of the CD function, namely that its elasticity of substitution equals one on every point of the isoquant map,[205]

$$y = A x_1^{\alpha_1} x_2^{\alpha_2} e^{\gamma\left(\frac{x_1}{x_2}\right)} \tag{4.26}$$

The inclusion of the x_1 / x_2 term results in a more realistic treatment for the elasticity of substitution by allowing it to vary according to the ratio of the two inputs. However, the inputs Ferguson referenced were aggregations of capital and labor, thus the best use of this function is for macroeconomic studies.

Nerlove-Ringstad Production Function

In 1963 Marc Nerlove studied the returns to scale of the electrical industry of the United States.[206] He presupposed that the underlying production function was of the Cobb-Douglas type without fully specifying that function. Upon the data available he did, however, propose two

---

[205] C. E. Ferguson, "Capital-Labor Substitution and Technological Progress in the U.S.: Statistical Evidence from a Transcendental Production Function" (Memo, 1965).

[206] Marc Nerlove, "Returns to Scale in Electricity Supply," in *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld,* ed. Carl Christ (Standford, California: Stanford University Press, 1963).

different cost functions.[207] In 1967 Vidar Ringstad derived the underlying production functions which take the forms,[208]

$$y = \left(A \prod_{i=1}^{n} x_i^{\alpha_i} \right)^{\frac{1}{\gamma + \beta \ln y}} \tag{4.27}$$

and

$$y = Ae^{\left(-a \prod_{i=1}^{n} x_i^{-\alpha_i} \right)} \tag{4.28}$$

Both equations (4.27) and (4.28) are non-homogeneous functions with variable returns to scale. The extension to *n*-inputs lends their use to microeconomic analysis although originally used in macroeconomic analysis.

Sato-CES Production Function

A new approach to factor substitution arose in 1967 with the Sato CES function.[209] Previously developed constant elasticity functions treated all pair-wise elasticities as equal. The treatment in the Sato CES function assumes that factors of production may logically be separated into subsets. These subsets have constant and equal elasticities among the factors constituting the subset, but each subset may exhibit a different elasticity. Normally the production function is expressed as $y = F(x) = F(x_1, x_2, \cdots, x_n)$. Under the condition of disaggregation of inputs and assuming that variable elasticities exist, the function can be constructed as a set of input bundles as $y = f(N_1, N_2, \cdots, N_s)$ such that each input bundle consists of one or more of the original input

[207] Ibid.
[208] Vidar Ringstad, "Econometric Analyses Based on a Production Function with Neutrally Variable Scale-Elasticity," *Swedish Journal of Economics* 69, no. 2 (1967): 117, 22.
[209] K Sato, "A Two-Level Constant-Elasticity-of-Substitution Production Function," *The Review of Economic Studies* 34, no. 2 (1967).

factors x_i. The production function can then be written as $y = f\left(g\left(x^{(1)}\right), g\left(x^{(2)}\right), \cdots g\left(x^{(n)}\right)\right)$. If we assume strong separability between the input bundles the function can be expressed in additive form as $y = f\left(g\left(x^{(1)}\right) + g\left(x^{(2)}\right) + \cdots + g\left(x^{(n)}\right)\right)$. Applying this formulation to the CES function we obtain,

$$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho}\right)^{-\frac{1}{\rho}} \mid Z_s = \left(\sum_{i \in s} \gamma_i^s \left(x_i^s\right)^{-\rho_s}\right)^{\frac{-1}{\rho_s}} \tag{4.29}$$

Equation (4.29) is a two-level CES function in that y is a CES function in z_s and z_s is a function of the $x_{i's}$ and the inter-class elasticity of substitution is constant. This formulation is homogeneous in degree one and therefore exhibits a fixed constant returns to scale. Concerned that equation (4.29) provided only for fixed constant returns to scale Sato introduced an additional parameter to give it a fixed variable return to scale as follows,

$$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho}\right)^{-\frac{v}{\rho}}. \tag{4.30}$$

In an attempt to correct the constant inter-class elasticity of substitution Sato introduced ρ_s' to allow for unequal inter-class elasticities as follows,

$$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho_s'}\right)^{-\frac{v}{\rho}}. \tag{4.31}$$

Equations (4.30) and (4.31) are homogeneous of degree one with (4.30) exhibiting equal inter-class elasticities and (4.31) exhibiting unequal elasticities. Despite the apparent unwieldiness of these functions they provide three variations of a two-level system of analysis that simultaneously evaluates the subsets and the global function. Developed for analysis of aggregated national supply of electricity, and thus a macroeconomic application, its extendibility to n-inputs make it suitable for microeconomic analysis, however, the complexity of the form make estimation problematic particularly as the number of inputs increases.

Revankar Generalized Production Function

In his 1967 doctoral dissertation Nagesh Revankar presented a new production function that corrects for the constant elasticity of substitution and fixed returns to scale characteristics of the previous production functions.[210] The resulting formulation is,

$$y = Ax_1^{(1-\alpha\gamma)}\left(x_2 + (\gamma - 1)x_1\right)^{\alpha\gamma}, \tag{4.32}$$

where x_1 and x_2 are capital and labor inputs and α, β and γ are parameters to be estimated. Thus the elasticity of substitution varies with the ratio of capital and labor. Revankar proceeded to further generalize the VES production function to account for variable returns to scale with,

$$y = Ax_1^{v(1-\alpha\gamma)}\left(x_2 + (\gamma - 1)x_1\right)^{v\alpha\gamma}, \tag{4.33}$$

thus converting it into a non-homogeneous function of degree v. The Revankar production function generalizes the elasticity of substitution and returns to scale factors and is thus a more realistic representation. However, its aggregation of capital and labor does not lend itself to extension to multiple inputs and therefore is unsuited for microeconomic analysis.

Kmenta Production Function

In 1967 Jan Kmenta wrote *On the Estimation of the CES Production Function*[211] in which he notes that restriction of the ACMS CES production function to a constant return to scale. In his article he proposed a method of estimation for the more general case of variable returns to

---

[210]Nagesh S. Revankar, "Production Functions with Variable Elasticity of Substitution and Variable Returns to Scale" (Dissertation, University of Wisconsin, 1967).
[211] Jan Kmenta, "On Estimation of the CES Production Function," *International Economic Review* 8, no. 2 (1967).

scale.[212] Kmenta demonstrates that an approximation of the ACMS production function, equation(4.19), can be achieved first by a logarithmic transformation of the ACMS function,

$$\ln y = \ln A - \nu/\alpha \ln\left(\beta x_1^{-\alpha} + (1-\beta)x_2^{-\alpha}\right), \tag{4.34}$$

then by a Taylor expansion around the term $\alpha = 0$ resulting in,

$$\ln y = \ln - \nu\beta \ln x_1 + \nu(1-\beta)\ln x_2 - 1/2\,\alpha\nu\beta(1-\beta)(\ln x_1 - \ln x_2)^2. \tag{4.35}$$

Equation (4.35) can be estimated using OLS methodology. The underlying production function is,

$$y = ae^{-.5(1-\alpha)\alpha\gamma\nu(\ln x_1 - \ln x_2)^2 + \alpha\nu\ln x_1 + (1-\alpha)\nu\ln x_2}. \tag{4.36}$$

Despite the complexity of this formulation the estimation by ordinary least squares is straight forward by means of a logarithmic transformation (precisely what Kmenta achieved). The function is homogenous in degree v thus exhibiting variable returns to scale. However, this function is not suitable for multiple input microeconomic analyses.

In 1971, in *Economies of Scale and the Form of the Production Function,*[213] Zvi Griliches and Vidar Ringstad, in a study of manufacturing in Norway, produced a nearly identical production function as that of Kmenta's. Using the same Taylor series expansion they reformulated Kmenta's function as,

$$\ln\left(y/x_1\right) = \delta_0 + \delta_1 \ln x_1 + \delta_2 \ln\left(x_2/x_1\right) + \delta_3\left(\ln\left(x_2/x_1\right)\right)^2. \tag{4.37}$$

Equations (4.35) and (4.37) are non-homogenous functions exhibiting variable returns to scale and can be estimated using OLS. Both equations are used extensively in macroeconomic

---

[212] Ibid., 180.

[213] Zvi Griliches and Vidar Ringstad, *Economies of Scale and the Form of the Production Function: An Econometric Study of Norwegian Manufacturing Establishment Data*, ed. J. Johnston, D. W. Jorgenson, and J. Waelbroeck, vol. 72, Contributions to Economic Analysis (Amsterdam, The Netherlands: North-Holland Publishing Company, 1971).

analysis but are not suitable for *n*-input microeconomic analysis. The importance of these equations may well rest in their being the foundation of the Sargan's Log-Quadratic production function discussed below.

Bruno Constant Marginal Share (CMS) Production Function

In 1968, in an attempt to generalize the Cobb-Douglas production function under condition of labor and capital market disequilibrium Michael Bruno developed the following production function,[214]

$$y = Ax_1^{\alpha} x_2^{1-\alpha} - mx_1 . \tag{4.38}$$

This formulation directly addresses the issue of the elasticity of substitution in a growing economy where prices of capital and labor vary with aggregated levels of national output and thus is strongly macroeconomic in character. Application of the CMS production function in microeconomic analysis is limited to the 2-input case.

Lu –Fletcher VES Production Function

In 1968 Yao-chi Lu and Lehman B. Fletcher introduce the production function bearing their names.[215] In their article they argue that the CES production function is limiting in that it assumes a constant elasticity of substitution along an isoquant, though not necessarily of unity. Lu and Fletcher show that the value of the elasticity of substitution is sensitive to changes in the capital to labor ratio. The production function they developed is as follows,

---

[214] Michael Bruno, "Estimation of Factor Contribution to Growth Under Structural Disequilibrium," *International Economic Review* 9, no. 1 (1968).

[215] Yao-chi Lu and Lehman B. Fletcher, "A Generalization of the CES Production Function," *The Review of Economics and Statistics* 50, no. 4 (1968).

$$y = A\left(\beta x_1^{-\rho} + (1-\beta)\gamma\left(\frac{x_1}{x_2}\right)^{\delta(1-\rho)} x_2^{-\rho}\right)^{-\frac{1}{\rho}}. \tag{4.39}$$

Lu and Fletcher point out that their production function exhibits all the normal characteristics required of production functions. In particular this function is homogenous in degree one and now exhibits variable elasticities of substitution along an isoquant. In using only capital and labor as functional inputs, and in particular the use of the ratio of capital to labor, this function is, however, highly macroeconomic in application. The authors further point out that this function is very difficult to extend to more than two inputs and that because the function is nonlinear in parameters it does not yield easily to estimation. Their function does present a case of generalizing an earlier form in that when $\delta = 0$ it forces γ to go to one and thus the function reverts to the original CES production function.

Sato-Hoffman Family of VES Production Functions

Ryuzo Sato and Ronald F. Hoffman offered in their 1968 article *Production Functions with Variable Elasticity of Factor Substitution: Some Analysis and Testing*, three (VES) production functions.[216] They found motivation for their work by challenging the prevailing notion of the constancy of the elasticity of substitution with the claim that "Once one drops the assumption of constancy and admits a variable elasticity of factor substitution (VES), the resultant production function depends on the assumptions involved in the elasticity of factor substitution function."[217] The first VES production function relies upon the condition that the

[216] Ryuzo Sato and Ronald F. Hoffman, "Production Functons with Variable Elasticity of Factor Substitution: Some Analysis and Testing," *The Review of Economics and Statistics* 50, no. 4 (1968).
[217] Ibid., 453.

elasticity of substitution is a function of the capital-labor ratio $k = {}^{x_2}\!/_{x_1}$, where x_2 is capital and x_1 is labor such that $\alpha(k) = ak + b$. The resulting production function is given by,

$$y = A\left(\frac{x_2}{x_1}\right)^{\beta} e^{\alpha\left({}^{x_2}\!/_{x_1}\right)}. \tag{4.40}$$

This function is variable in elasticity and exhibits constant returns to scale. The second VES production function is based upon σ being a linear function of k such that $\sigma = a + bk$. The resulting production function is given by,

$$y = Ax_2^{\alpha\left({}^{1}\!/_{1+\gamma}\right)}\left(x_1 + \left({}^{\beta x_2}\!/_{1+\gamma}\right)x_1\right)^{\alpha\left({}^{\gamma}\!/_{1+\gamma}\right)}. \tag{4.41}$$

Equation (4.41) exhibits variable elasticity of substitution and variable returns to scale formulation, with the α term representing the returns to scale factor. The last production function derived is based upon σ as a function of time such that $\sigma = \alpha + \beta t$ and is given by,

$$y = A\left(\delta x_1^{\sigma-\left({}^{1}\!/_{\sigma}\right)} + (1-\delta)\,x_2^{\sigma-\left({}^{1}\!/_{\sigma}\right)}\right)^{{}^{\sigma}\!/_{(\sigma-1)}}. \tag{4.42}$$

This formulation exhibits variable elasticity of substitution and constant returns to scale. In these three production functions Sato and Hoffman demonstrated that the elasticity of substitution along an isoquant is variable under certain conditions. Each of these production functions are strongly macroeconomic in nature due to their development being based upon aggregated national values of capital and labor.

Sato's CEDD Production Functions

Ryuzo Sato developed four new production functions in 1970 based upon the concept of 'constant elasticity of derived demand'. These production functions appeared in *The Estimation* *of Biased Technical Progress and the Production Function.*[218] The following production functions were derived from the basic position that the elasticity of substitution σ equals the product of the share of labor (share of capital) and the elasticity of derived demand for capital per unit of labor (for labor per unit of capital input) such that $\sigma_l = \beta E_k$ and $\sigma_k = \alpha E_l$.[219] Based upon these relationships and by assuming σ_l and $\sigma_k \neq 1$, and allowing x_1 to represent labor and x_2 to represent capital, the following two production functions were derived,[220]

$$y = \left(\alpha_1 / 1 - \delta_1 \right) x_1^{\delta_1} x_2^{(1-\delta_1)} + \beta_1 x_1 \qquad (4.43)$$

and

$$y = \left(\alpha_2 / 1 - \delta_2 \right) x_2^{\delta_2} x_1^{(1-\delta_2)} + \beta_2 x_2 . \qquad (4.44)$$

Equations (4.43) and (4.44) exhibit variable elasticities of substitution and constant return to scale. Both are strongly macroeconomic in character. When the assumption about σ_l and σ_k is changed to make both equal to one then the following two production functions are obtained,[221]

$$y = \alpha_1 x_1 \ln\left(\frac{x_2}{x_1} \right) + \beta_1 x_1 \qquad (4.45)$$

and

$$y = \alpha_2 x_2 \ln\left(\frac{x_1}{x_2} \right) + \beta_2 x_2 . \qquad (4.46)$$

Equations (4.45) and (4.46) exhibit variable elasticities of substitution and variable return to scale. Both are strongly macroeconomic in character.

---

[218] Ryuzo Sato, "The Estimation of Biased Technical Progress and the Production Function," *International Economic Review* 11, no. 2 (1970).
[219] Ibid., 188.
[220] Ibid., 189.
[221] Ibid., 190.

Chu, Aigner, and Frankel Log-Quadratic Production Function

In 1970, in *On the Log-Quadratic Law of Production*, S. F. Chu, D. J. Aigner and M. Frankel proposed a new production function designed to be non-homogeneous and exhibiting varying returns to scale and variable elasticity of substitution.[222] Their production function is given by,

$$y = A\left(x_1 / \overline{x}_1 \right)^{\alpha_1\left((1-\ln x_1)/\ln \overline{x}_1 \right)} \left(x_2 / \overline{x}_2 \right)^{\alpha_2\left((1-\ln x_2)/\ln \overline{x}_2 \right)}, \tag{4.47}$$

where A, α_1, α_2, $\overline{x}_1$, and $\overline{x}_2$ are parameters to be estimated. This formulation exhibits both variable elasticities of substitution and variable returns to scale. It is also highly macroeconomic in nature.

Vazquez VES Production Functions

Appearing in *Homogeneous Production Functions with Constant or Variable Elasticity of Substitution* Andres Vazquez proposed three new VES production functions in 1971.[223] His motivation was to develop new algebraic forms of the CES production function but with variable elasticities of substitution. In order that he might obtain variable elasticities he relies upon the basic definition of homogeneous functions and rewrites the standard formula for σ by introducing the degree of homogeneity term λ to obtain $\sigma = \dfrac{f'(\lambda f - xf')}{(\lambda - 1)xf'^2 - \lambda x f f''}$.[224] In the first VES production function σ is conditioned by assuming a linear relationship between the

---

[222] S. F. Chu, D. J. Aigner, and M. Frankel, "On the Log-Quadratic Law of Production," *Southern Economic Journal* 37, no. 1 (1970): 32.

[223] Andres Vazquez, "Homogeneous Production Functions with Constant or Variable Elasticity of Substitution," *Journal of Institutional and Theoretical Economics* 127, no. 1 (1971).

[224] Ibid., 13.

average product of labor, the marginal production of labor and the capital-labor ratio exist such that $y/x_1 = \alpha + \beta \frac{dy}{dx_1} + \gamma \frac{x_2}{x_1}$, where x_1 and x_2 represent labor and capital respectively. The resulting production function is given by,

$$y = X_2^{\left(1-1/\beta\right)} X_1^{\left(1/\beta\right)} + \left(\alpha/1-\beta\right) X_1 + \gamma X_2 . \tag{4.48}$$

Equation (4.48) reduces to the Cobb-Douglas function when $\alpha = 0$ and $\gamma = 0$ and to Bruno's CMS function when $\gamma = 0$. In the special case when $\alpha = 0$ then Equation (4.48) reduces to,

$$y = X_2^{\left(1-1/\beta\right)} X_1^{\left(1/\beta\right)} + \gamma X_2 \tag{4.49}$$

Assuming that the ratio of the elasticities of the inputs and the capital-labor ratio are of the form $\frac{E_{x_2}}{E_{x_1}} = \alpha + \beta \left(\frac{x_2}{x_1} \right)^{\gamma}$ Vazquez formulated his third production function as,

$$y = A\left(\beta x_1^{-\gamma/(1+\alpha)} + (1+\alpha)\left(\frac{x_2}{x_1} \right)^{-\alpha\gamma} x_2^{-\gamma/(1+\alpha)} \right)^{-\lambda/\gamma(1+\alpha)} . \tag{4.50}$$

Whereas the first functions exhibited both variable elasticities of substitution and constant return to scale, this third function exhibits VES and VRTS. All three functions are strongly macroeconomic in character.

Sargan Production Function

With the publication of *Qualified Manpower and Economic Performance*[225] in 1971 the log-quadratic production function was popularized. As part of a team studying the electrical engineering industry J. Denis Sargan used the following production function,

$$\ln y = \ln A + \sum_{i=1}^{n} \alpha_i \ln x_i + \tfrac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j \,. \tag{4.51}$$

By inspection of this equation we note it permits *n*-inputs and exhibits both VES and VRTS characteristics.[226] Although originally used as a macroeconomic tool to study the electrical engineering industry in England its application to microeconomic analysis is obvious. The rapidly expanding number of parameter that must be estimated as the number of inputs increase makes estimation extremely problematic.

Christensen, Jorgenson and Lau (CJL) Transcendental Logarithmic Production Function

The transcendental logarithmic production function, commonly referred to by the term ‘translog’, first appeared in 1971, in a short two page note titled *Conjugate Duality and the Transcendental Logarithmic Production Function.*[227] A more thorough and detailed treatment occurred in 1973 in *The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-68.*[228] The CJL translog production function and Sargan’s production function are identical (see Equation(4.51)). While each was developed independently of the other the translog nomenclature has received the most wide spread acknowledgement. In

---

[225] Richard Layard et al., *Qualified Manpower and Economic Performance: An Inter-plant Study in the Electrical Engineering Industry* (London, England: Allen Lane The Penguin Press, 1971).
[226] Ibid., 154-55.
[227] Laurits R. Christensen, Dale W. Jorgenson, and Lawrence J. Lau, "Conjugate Dualtiy and the Transcendental Logarithmic Production Function," *Econometrica* 39, no. 4 (1971).
[228] Ernst R. Berndt and Laurits R. Christensen, "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-68," *Journal of Econometrics* 1, no. 1 (1973).

The Translog Production Function: Some Evidence from Establishment Data[229] Vittorio Corbo notes the advantages of the translog production function. Corbo begins by noting that the two most commonly used production functions – the Cobb-Douglas and CES production functions – impose restrictions on the properties of the underlying technology. In the case of the Cobb-Douglas production function it restricts all partial elasticities of substitution to equal one, and in the case of the CES production function it restricts the elasticities of substitution to be constant and equal (although not necessarily equal to one) for any pair of inputs for all points in input space; the translog production function does not impose these restrictions.[230]

The translog production function is a non-homogeneous function in its general form. If the empirical researcher finds that the $\beta_{ij's}$ are not significantly different from zero the general form reduces to the Cobb-Douglas form which is homogenous in the degree of the sum of its exponents. Also the translog exhibits variable elasticities of substitution. Although initially used in the study of aggregated inputs in various industries its expansion to *n*-inputs makes it useful for microeconomic analysis, however, the rapidly increasing number of parameters that must be estimated as the number of inputs increase make it highly problematic particularly with small data sets.

Diewert Generalized Leontief Production Function

Utilizing the Shepard Duality Theorem W. Erwin Diewert presented in *An Application of the Shepard Duality Theorem: A Generalized Leontief Production Function*[231] a production function that generalized the two-input ACMS production function to *n*-inputs. Diewert notes

[229] Vittorio Corbo and Patricio Meller, "The Translog Production Function," *Journal of Econometrics* 10, no. 2 (1979).
[230] Ibid., 193-94.
[231] W. Erwin Diewert, "An Application of the Shepard Duality Theorem: A Generalized Leontief Production Function," *The Journal of Political Economy* 79, no. 3 (1971).

that when we ask for a production function that exhibits an arbitrary constant elasticity of substitution that such a function only exists for the two-input case which has been demonstrated as the ACMS production function discussed early in this dissertation. Citing the impossibility theorem he states that it is not possible to extend a production function of arbitrary constant elasticity to the *n*-input case. He takes as the task of his paper to show that a production function of *n*-inputs with non-constant elasticities between pair-wise inputs is possible. The form he developed is,

$$y = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} x_i^{1/2} x_j^{1/2} . \tag{4.52}$$

The generalized Leontief production function exhibits variable elasticities of substitution and constant returns to scale. The expansion to accommodate *n*-inputs makes this production function particularly useful in microeconomic analysis with the caveat that the increasing number of parameters accompanying an increase in the number of inputs can make the estimating process problematic particularly with small data sets. When the $\alpha_{ij's} = 0$ for $i \neq j$ this function reduces to a linear production function.

Kadiyala Production Function

In 1972 Kadiyala presented a paper *Production Functions and Elasticity of Substitution*[232] wherein he developed a production function in which the elasticity of substitution varied along an isoquant. For the two-input case the production function takes the form,

$$y = A\left(\alpha_{11} x_1^{2\beta_0} + 2\alpha_{12} x_1^{\beta_1} x_2^{\beta_2} + \alpha_{22} x_2^{2\beta_0}\right). \tag{4.53}$$

---

[232] Koteswara Rao Kadiyala, "Production Functions and Elasticity of Substitution," *Southern Economic Journal* 38, no. 3 (1972).

He notes that the function is homogenous and that the elasticity of substitution varies with the input ratio along the isoquant from $\sigma = 0$ on one end to $\sigma = \infty$ on the other end.[233] Additionally this form reduces to the CES function when $\alpha_{12} = 0$, the Lu-Fletcher VES production function when $\alpha_{22} = 0$, and the Sato-Hoffman production function when $\alpha_{11} = 0$.[234] The Kadiyala form contains within it the Cobb-Douglas form, Leontief form, and linear form under conditions. When $\beta_0 = \beta_1 = \beta_2 \to 0$ then the Kadiyala form reduces to the Cobb-Douglas form; when $\beta_0 = \beta_1 = \beta_2 \to -\infty$ it reduces to the Leontief form, and when $\beta_0 = \beta_1 = \beta_2 \to 1/2$ and $\alpha_{12} \to 0$ it becomes a linear form.[235] Kadiyala concludes his paper by noting that the two-input case can be extended to *n*-inputs as follows,

$$y = A\left(\sum_{i=1}^{n} \alpha_{ii} x_i^{2\beta} + 2\sum_{i<j}^{n} \alpha_{ij} x_i^{\beta} x_j^{\beta}\right), \tag{4.54}$$

such that $\alpha_{ij} \geq 0,\ i,\ j = 1, \cdots, n$ and $\sum_{i=1}^{n} \alpha_{ii} + 2\sum_{i<j}^{n} \alpha_{ij} = 1$. Whereas Equation (4.53) is strongly macroeconomic in character the expansion to *n*-inputs in Equation (4.54) permits both macro and microeconomic analysis with the caveat that the increasing number of parameters accompanying an increase in the number of inputs can make the estimating process problematic particularly with small data sets.

Vinod Production Function

In a study of the telecommunications industry for Bell Labs in 1972 Hrishikesh Vinod presented a new non-homogeneous production as found in *Nonhomogeneous Production*

---

[233] Ibid., 283.
[234] Ibid., 282.
[235] Ibid., 283.

Functions and Applications to Telecommunications.[236] This production function is a modification of the Cobb-Douglas form allowing for an interaction term by the product of the logarithms of inputs as follows,

$$y = e^{\alpha_0} x_1^{\alpha_1 + \alpha_3 \ln x_2} x_2^{\alpha_2} , \tag{4.55}$$

which can also be expressed,

$$\ln y = \alpha_0 + \alpha_1 \ln x_1 + \alpha_2 \ln x_2 + \alpha_3 \ln x_1 \ln x_2 .^{237} \tag{4.56}$$

Vinod notes that this expression is both VES and VRTS plus the form is linear in parameters thus lending itself to estimation with OLS. The Vinod production function can be extended to *n*-inputs as follows,

$$\ln y = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + \sum_{i<j}^{n} \alpha_{ij} \ln x_i \ln x_j . \tag{4.57}$$

Both equations were developed for macroeconomic analysis of the telecommunications industry, however, the extension of Equation (4.57) to the *n*-input case makes it appropriate for microeconomic analysis as well, once again with the caveat that the increasing number of parameters accompanying an increase in the number of inputs can make the estimating process problematic particularly with small data sets.

Lovell Production Functions

Concerned about perceived shortcomings of the CES and VES formulations popular at the time C. A. Knox Lovell proposed, in 1973, one CES and two VES production functions in

---

[236] Hrishikesh D. Vinod, "Nonhomogeneous Production Functions and Applications to Telecommunications," *The Bell Journal of Economics and Management Science* 3, no. 2 (1972).
[237] Ibid., 532.

Estimation and Prediction with CES and VES Production Functions.[238] Lovell's attempt at 'correcting' the earlier formulation were based in a desire to estimate or evaluate all the dimensions of the production function including the possibility of technological change over time. The first production function developed was of the CES variety with a technological change element added as follows,

$$y = Ae^{\lambda t}\left(\delta x_1^{-\rho} + (1-\delta)x_2^{-\rho}\right)^{-1/\rho}, \tag{4.58}$$

where λ measures technological change over time t. Minus the $e^{\lambda t}$ term we have the familiar ACMS CES formulation. The second production function is of the VES variety generated from the side condition that $\sigma = \alpha + \beta k$ and takes the form,

$$y = Ae^{\lambda t}\left((1+\beta)x_1 x_2^{\beta} + \alpha x_2^{1+\beta}\right)^{1/1+\beta}. \tag{4.59}$$

His second VES formulation began with the side condition that $\sigma = k\left(\frac{1}{\alpha+\beta k} - 1\right)$ and takes the form,

$$y = Ae^{\lambda t} x_1^{\alpha} x_2^{1-\alpha} e^{\beta k} \tag{4.60}$$

All three production functions are fixed returns to scale with the CES function having constant elasticities of substitution while the VES functions exhibit variable elasticities along the isoquant. Each production function is highly macroeconomic in nature and thus not suitable for microeconomic analysis requiring multiple inputs.

[238] C. A. Knox Lovell, "Estimation and Prediction with CES and VES Production Functions," *International Economic Review* 14, no. 3 (1973).

Diewert "Generalized Cobb-Douglas" Production Function

In 1973 W. Erwin Diewert demonstrated a generalization of the Cobb-Douglas production function in *Separability and a Generalization of the Cobb-Douglas Cost, Production and Indirect Utility Functions*.[239] The objective was a production function based upon the Cobb-Douglas formulation that made no assumptions regarding the constancy of the partial elasticities of the input factors except that they should sum to one thus ensuring a constant return to scale formulation. The resulting formulation is,[240]

$$y = A\prod_{i=1}^{n}\prod_{j=1}^{n}\left(\tfrac{1}{2}x_i + \tfrac{1}{2}x_j\right)^{\alpha_{ij}} , \tag{4.61}$$

such that $\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_{ij} = 1$. This condition ensures that the function is a constant return to scale function. Given a logarithmic transformation of Equation (4.61) it is possible to estimate the $\alpha_{ij's}$. With the aid of modern computer based statistical models it is possible to set the side condition that $\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_{ij} = 1$ and estimate the function. However, it may be more important to run such a model without the side condition to determine if any of the $\alpha_{ij's}$ are less than zero and whether the $\alpha_{ij's}$ sum to one. If any of the $\alpha_{ij's}$ are less than zero it may be necessary to test of concavity of the function and convexity of the isoquants. If the $\alpha_{ij's}$ do not sum to one then the function may be of the variable returns to scale type. Overall we may say that this type of analysis gives us more information about the nature of the underlying production function.

Under the initial conditions set for the function it is a homogenous function of degree one and therefore constant return to scale formulation. The extension to *n*-inputs makes it

[239] W. Erwin Diewert, "Separability and a Generalization of the Cobb-Douglas Cost, Production and Indirect Utility Functions," (Vancouver, British Columbia, Canada: Stanford University and University of British Columbia, 1973).
[240] Ibid., 21.

particularly appropriate for microeconomic analysis but once again with the caveat that the increasing number of parameters accompanying an increase in the number of inputs can make the estimating process problematic particularly with small data sets.

Denny Generalized Quadratic Production Function

In 1974 Michael Denny wrote *The Relationship Between Functional Forms for the Production System*[241] with the intent to demonstrate a new production function he termed the 'generalized quadratic' production function. His motivation was the desire to create a formulation that linked previous production functions in a way that permits empirical researchers to test for any of the included forms namely Diewert's generalized Leontief, the CES, and Cobb-Douglas forms.[242] The necessity of this work, he felt, arose from the difficulty researchers found in selecting specific forms of the production function in their work.[243] The form he developed is as follows,[244]

$$y = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} x_i^{\beta\gamma} x_j^{\beta(1-\gamma)} \right)^{\mu/\beta}. \tag{4.62}$$

Setting $\gamma = \frac{1}{2}$ and $\alpha_{ij} = 0$ for all $\alpha_{i \neq j}$ reduces Equation (4.62) to the Uzawa-McFadden CES production function, Equation(4.21). Setting $\gamma = \frac{1}{2}$, $\alpha_{ij} = 0$ for all $\alpha_{i \neq j}$ and as μ and $\beta \rightarrow 0$ Equation (4.62) reduces to the Cobb-Douglas production function, Equation(4.2). The Diewert

[241] Michael G. S. Denny, "The Relationship Between Functional Forms for the Production System," *The Canadian Journal of Economics* 7, no. 1 (1974).
[242] Ibid., 21.
[243] Ibid., 22.
[244] Ibid., 29.

production function, Equation(4.52), is obtained when $\gamma = \frac{1}{2}$ and $\beta = 1$.[245] The capabilities of modern software modeling and statistical applications make the estimation of Equation (4.62) more straight forward than would have been possible when this article first appeared. Conducting such an analysis would easily permit a researcher to determine which form may best represent the case under study.

The flexibility of the generalized quadratic is most evident when considering that, depending upon the value the estimated parameters take, the function may be homogeneous or nonhomogeneous as well as permitting constant or variable elasticities of substitution and variable returns to scale. The *n*-input nature of the function makes it excellent for microeconomic analysis with the same caveat as before that the increasing number of parameters accompanying an increase in the number of inputs can make the estimating process problematic particularly with small data sets.

Helmy Generalized Transcendental Production Functions

In 1981 in his doctoral dissertation *A Family of Transcendental Production Functions*[246] Aly Helmy derived three productions based upon the HCH transcendental production and one based upon CJL trans-log function. His aim was to develop transcendental production functions that allowed specifically for a factor of interaction between inputs. He referred to the first three production functions as 'generalized transcendental' production functions and the fourth as the 'quadratic logarithmic' production function. GTPF(1) function is expressed as,

[245] Ibid., 24. Each of the three preceding equations are derived here. It is not clear why Denny chose to exclude the efficiency factor A from his production function for its inclusion would not change his other calculations.
[246] Helmy, "A Family of Generalized Transcendental Production Functions."

$$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} x_i x_j}. \tag{4.63}$$

GTPF(2) is expressed,

$$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} \ln x_i \ln x_j}. \tag{4.64}$$

GTPF(3) is expressed,

$$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} x_i^{1/2} x_j^{1/2}}. \tag{4.65}$$

Equations (4.63) , (4.64) and (4.65) must also satisfy the restriction that $\beta_{ij} = \beta_{ji} = 0$ for $i = j$, i and $j = 1$ to n. We may further specify that when $\beta_{ij} = 0$ for all β_{ij} that these equations reduce to the transcendental production function. In turn if we add the additional specification that when $\beta_i = 0$ for all β_i these functions further reduce to the Cobb-Douglas production function.

The Quadratic Logarithmic production function is expressed,

$$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i (\ln x_i)^2}. \tag{4.66}$$

Equation (4.66) reduces to the Cobb-Douglas production function when $\beta_i = 0$ for all β_i. All of these production functions exhibit variable elasticities of substitution and variable returns to scale. Additionally the three GTPF's allow for interaction between the inputs. Statistical estimation of all of the fore mentioned Helmy production functions is problematic due to the large number of parameters that must be estimated and is particularly troublesome with small data sets.

Summary

In this chapter 53 production functions were enumerated and their salient characteristics noted. Appendix G provides a recapitulation of these production functions and their characteristics. The task next turns to determining which of these production functions is suitable for empirical study in the context of architectural firms; that is the subject of Chapter 5, Choice of Functional Form.

CHAPTER 5

CHOICE OF FUNCTIONAL FORM

If all the economists were laid end to
end, they would not reach a conclusion.
George Bernard Shaw[247]

Introduction

The survey of production functions presented in chapter 4 enumerated and described the properties of 53 mathematical forms of the production function. Such a large number of forms would prove daunting to an empirical researcher faced with estimating each function and all the associated parameters. At the end of such a task the researcher certainly would feel that much time and effort had been wasted when it turned out that many of the candidate forms proved either inappropriate to the research at hand or proved statistically problematic. Clearly a more focused smaller list of functions is called for. It is to the winnowing of that list to a more manageable number for empirical work that this chapter is dedicated. The first part of this chapter develops a winnowing methodology by which choice among alternatives may be made. The second part of this chapter applies that methodology and provides the final list of production function forms that will be used in chapter 6 to demonstrate the application of production theory to architectural firms.

Methodology

Introduction

The first part of this chapter formulates the criteria for selection (or elimination) of mathematical forms of production functions. The broad outline of this methodology is laid out

[247] Michael Moncur, "The Quotation Page," www.quotationspage.com/quotes/George_Bernard_Shaw.

by Melvyn Fuss, Daniel McFadden and Yair Mundlak in *Production Economics: A Dual Approach to Theory and Applications*.[248] Broadly speaking 6 criteria are proposed. These criteria are 1) application and objective, 2) minimization of maintained hypotheses, 3) parsimony in parameters, 4) ease of interpretation, 5) computation ease, and 6) interpolative and extrapolative robustness.

Application and Objective

As noted throughout chapter 4 the motivation for the construction of a specific production function stemmed from a variety of concerns or proposed uses for those forms. Therefore it should not be surprising that the intended application or research objective should favor one or more forms while forming the basis for elimination of others. We may distinguish among the various applications of production functions and the intended purpose or objective of empirical research utilizing these forms by placing them into two categories.[249] The first categorization situates the application or objective along the continuum of economic theory from greatest aggregation to least, that is, the distinction between macroeconomics as the study of large systems (national economies, industries, and sectors) against microeconomics as the study of individual and firms. By stating that the objective in this study is the identification of production functions appropriate for the study of individual architectural firms and in particular those allowing *n*-inputs, where $n > 2$, then those production functions developed for macroeconomic analysis and not modified for use in microeconomic analysis may now be eliminated. This alone eliminates 27 of the 53 forms surveyed in chapter 4.

[248] Fuss and McFadden, *Production Economics: A Dual Approach to Theory and Applications, Vol I and II*: 219-25.
[249] Ibid., 220.

The second categorization distinguishes between those production functions useful in analytical studies and those used for predictions.[250] The first group includes studies of the technical attributes of the production function in the context of a specific data set. Such technical attributes include constancy of returns to scale, the constancy of the elasticity of substitution, and others. The second group is used in a predictive or optimization context. Examples of this second group include predicting the level of use of certain resources such as energy or the optimization of the allocation of resources, which is the objective this study is most closely associated. The difficulty one encounters in this second categorization is that not all production functions fall neatly in to one group and not the other. This is particularly true with the group of production functions previous identified as generalized functions. The generalized or nested functions are those functions which subsume other functions such that when estimated, one or more of their parameters approach certain extreme values (generally zero, one, or infinity) and take on the form of the lesser included model. These functions may be used to test specific technical conditions but are also useful in distinguishing simpler forms from their more complex parent. In the context of functional forms falling into this second general categorization some subjective judgment is applied in determining those of high potential and those with low potential as pertains to the objective of this study.

Minimization of Maintained Hypotheses

Chapter 3 presents a discussion of the properties of production functions that economists believe are true hypotheses about the production function itself or the economic conditions in which they arise.[251] These hypotheses are not generally tested as part of an economic analysis

[250] Ibid.
[251] Ibid., 222-24.

but are held to be true none-the-less. These hypotheses may be grouped into four categories. The first category deals with the basic axioms of the nature of the technology. Included in this group is whether the input set is closed, whether positive output exists in the absence of all inputs under study, whether output requires the presence of all inputs, etc. The second group addresses hypotheses of technological and behavioral assumptions. These are assumptions are not necessarily held widely as true but are credible in the context of the research at hand. Included in this group is whether the production technology is a convex technology (as opposed to the production function outputs being convex for which we can test), the constancy of input and output prices, or whether the entrepreneur in reality is a cost minimizer or profits maximizer. The third group of hypotheses enables the statistical analysis of data. These hypotheses are considered harmless approximations of reality and include that the errors are independent and normally distributed, variances are homoscedastic, and that intermediate inputs are separable from primary inputs. The final group of hypotheses relate to the parametric form of the production function. Here the analyst makes an assumption about the form that is accepted but not tested in their analysis. Such assumptions may involve technical aspects of the form in relation to the assumed economic conditions. One example would be assuming constant returns to scale without testing for variable returns. The general principle that flows from consideration of these various maintained hypotheses is "that one should not attempt to test a hypothesis in the presence of maintained hypotheses that have less commonly accepted validity."[252] In consideration of these hypotheses (many of which apply to all production functions) the characteristics of generalized production functions enjoy favor over those that *a priori* assume certain conditions exist such as constant returns.

[252] Ibid., 223.

Parsimony in Parameters

An increase in the number of parameters utilized in a function form brings with it the potential for problems in the statistical analysis of the form. Among these issues are the likelihood of multicollinearity, a loss of degrees of freedom, large standard errors, and the need for larger data sets to ensure statistically valid analysis. Avoiding these issues, or at least reducing them to a minimum, requires employing the general principle "(a) functional form should contain no more parameters than are necessary for consistency with the maintained hypotheses."[253] Clearly the more generalized functional forms have more parameters than their simpler cousins but have the value, once the parameters are estimated, of indicating which, if any, of the simpler forms best approximates the true form. These more complex forms raise the issue of the minimum size of the data set used in the estimation process. In light of data sets composed of relatively large numbers of inputs and relatively small number of observations, which we will encountered in chapter 6, another 21 forms can now be eliminated for exhibiting a prohibitively large number of parameters.

Ease of Interpretation

An increase in the complexity of a functional form is normally accompanied by an increase in the potential difficulty in interpreting the estimated model. Assessing the elasticity of substitution, the separability of inputs, and interaction between inputs are examples of such problems. The principle flowing from these concerns is "(I)t is better to choose functional forms in which the parameters have an intrinsic and intuitive economic interpretation, and in which functional structure is clear."[254] In this winnowing process a high rating (5) is used to denote

[253] Ibid., 224.
[254] Ibid.

ease of interpretation and low rating (1) used to denote difficulty in interpretation of functional form.

Computation Ease

The measure of computation ease is inversely related to the degree of complexity found in the statistical analysis process required to estimate a given functional form. Until the advent of the personal computer and sophisticated statistical software the most common and simplest method of estimating the parameters in a multiple regression was ordinary least squares or OLS. Using OLS requires that the parameters be linear. In some circumstances the functional form may be transformable from its original non-linear form into one that is linear in parameters. The Cobb-Douglas function is one such example. However many of the functional forms enumerated in chapter 4 require the more complex and sophisticated statistical software available (e.g. SPSS, R, etc.) to successfully estimate their parameters. The general principle employed is that of seeking a carefully considered balance between computational complexity and the thoroughness required in empirical analysis.[255] In the winnowing process employed in this study a high rating equals highly easy to estimate, generally employing OLS, while a low rating equals difficult to estimate, generally employing a version of generalized least squares and requiring the use of a computer and a powerful statistical software package.

Interpolative and Extrapolative Robustness

Interpolative robustness requires that a functional be well-behaved within the range of observed data while extrapolative robustness requires the functional form to be well-behaved outside of the range of observed data. In either case the functional form must be consistent with

[255] Ibid., 224-25.

the maintained hypotheses upon which it is based. For interpolative robustness, should the consistency be in doubt, a numerical confirmation of behavior may become necessary in which case the functional form should permit convenient computation of relevant technical aspects of the form.[256] Often certain hypotheses such as positive marginal products or convexity in isoquants are held true for a functional form. Should doubt over the behavior in question ever arise the empirical researcher desires a functional form that permits easy verification. Extrapolative robustness is not easily confirmable as by definition the behavior in question lies outside the range of observed data. However, in some circumstances it can be shown that certain technical aspects of functional forms are true over all positive values of inputs and outputs. Given the extensive body of peer reviewed literature regarding the functional forms found in chapter 4 and lacking any sense that one or more maintained hypotheses may be found untrue in any of these functional forms it is unlikely that interpolative or extrapolative robustness is an issue in this study and is not further considered.

Winnowing the List

In this section the list is winnowed from the original 53 functional forms to a selection of five functional forms. These selected forms are discussed and ranked below. Two of the six criteria account for the elimination of 48 of the original 53 candidate forms. The criteria of application and objective accounted for 27 of the 47 eliminations. Under this criterion the basis for elimination of each of these forms stems from their exclusive use in macroeconomic analysis. In that none of these forms permits more than two inputs their unsuitability for use in microeconomic analysis is apparent. A further 21 forms where eliminated by the criteria of

[256] Ibid., 225.

parsimony in parameters. In each case the numbers of parameters to be estimated by statistical methods were excessively large. The data sets of completed projects of architectural firms contain between 5-10 labor inputs and have 25 or fewer observations. In each of the forms eliminated under this criterion the number of parameters to be estimated exceeds the number of observations present. Analysis under such conditions is not possible. Appendix H (Choice of Functional Form) displays a chart of all 53 functional forms along with an evaluation of their usefulness in this study. The five remaining forms are listed below.

Table 1. Final Candidate Production Functions.

Name	Function	Evaluation Criteria					
		Application & Objective	Maintained Hypotheses	Parsimony in Parameters	Ease of Interpretation	Computational Ease	Total
Cobb-Douglas	$y = A\prod_{1}^{n} x_i^{\alpha_i}$	2	1	5	5	5	18
Leontief	$y = \min(\beta_1 x_1, \beta_2 x_2, \cdots, \beta_n x_n)$	1	1	5	5	5	17
Uzawa-McFadden	$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha}\right)^{-\frac{v}{\alpha}}$	3	3	4	5	1	16
Mukerji	$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha_i}\right)^{-\frac{1}{\alpha_0}}$	3	4	3	4	1	15
Nerlove-Ringstad (1)	$y = \left(A\prod_{i=1}^{n} x_i^{\alpha_i}\right)^{\frac{1}{\gamma+\beta\ln y}}$	3	3	4	1	1	12

Cobb-Douglas Production Function

The Cobb-Douglas production function was the first and is arguably the most popular production function ever developed.[257] The advantages of the CD form are its computational ease, parsimony in parameters and the ease by which the results can be interpreted. Estimating the parameters of the CD function are rather straight forward requiring only a logarithmic transformation of the basic equation which renders it linear-in-parameters and then application of the ordinary least squares (OLS) statistical estimation methodology. The low parsimony in parameters, where the total number of parameters to be estimated is $2n+1$, in conjunction with the use of OLS means that the CD function can be estimated using less sophisticated software such as Excel®. The relatively simple mathematics of the CD form makes its use in optimization schemes straight forward and relatively easy. The principle disadvantage of the CD form stems from the number and type of maintained hypotheses it exhibits. The CD form assumes a fixed elasticity of substitution – permitting no variable elasticity of substitution among variables – a fixed return to scale, that all inputs are necessary in the production process, and that when one input level is zero that output is likewise zero.

Leontief Production Function

The second most commonly encountered production function is the Leontief production function. Advantages of the Leontief form are that it may be estimated using the OLS methodology, low parsimony in parameters with the number of parameters equaling the number of inputs, and that the resulting parameters directly indicate the appropriate ratio of inputs, or recipe, inherent in the production process represented by this form. Disadvantages inherent

[257] Alan Arthur Walters, "Production and Cost Functions: An Econometric Survey," *Econometrica* 31, no. 1/2 (1963): 5.

include the fact that it admits no substitution between inputs, that it assumes a fixed constant return to scale, and that should any input be zero then the resulting output is also zero. The fixed constant return to scale makes the form unsuitable for profit maximization calculations. Similarly, cost minimization or output maximization optimization routines become trivial. Cost minimization is just the simple summation of the relevant input costs and their associated usage level represented as, $c = \sum_{i=1}^{n} w_i x_i$. Thus the cost minimization position is not sensitive to wage rates or changes in relative wage rates. In a similar manner output maximization is just a function of the amount of the least available input. The Leontief production function is most useful when the assignment of resources is known in advance to hold to a predetermined recipe.

Uzawa-McFadden Production Function

The Uzawa-McFadden production function enjoys an advantage over the other CES production functions in that it permits both variable returns to scale and *n*-inputs while maintaining the other assumptions of the ACMS form. The principal disadvantages rest in the complexity of the form, whose estimation requires highly sophisticated computer software programs such as SPSS, and the high degree of difficulty in its use in optimization routines.

Mukerji Production Function

The advantage that the Mukerji formulation holds over the previous CES forms, namely that of permitting variable elasticities of substitution between inputs, is offset by the disadvantages in complexity of estimation due largely to the increase in the number of parameters that must be estimated which is $2n+2$. As with the Uzawa-McFadden form its interpretation and subsequent application in optimization schemes is highly problematic.

Nerlove-Ringstad Production Functions

The Nerlove-Ringstad formulation bears considerable similarity to the CD form upon which it is based and retains many of the assumptions, except permitting variable returns to scale, inherent in the CD form. As with the two preceding forms the number of parameters and the mathematical form make interpretation, estimation, and application difficult.

Summary

The winnowing process has reduced the list of candidate forms from 53 to the final 5 noted above. In chapter 6 each of these forms are further analyzed utilizing data obtained from industry in an attempt to find the best production function(s) for use in empirical research in the field of architecture.

CHAPTER 6

CASE STUDY

If you torture the data long enough it will confess.
Ronald Coase[258]

Introduction

This chapter presents a case study of a medium sized successful architectural firm located in the mid-west. The objectives of the case study are twofold. The primary purpose of this case study is the identification of those production functions, from the five candidates, that best represent the technologies employed by the firm and provides the soundest statistical basis for empirical study of other firms. A secondary objective is the identification of issues relevant to the statistical estimation of these production functions with the aim of improving subsequent empirical analyses.

The firm selected for this study is an architectural firm of long standing located in the mid-west. Its client base is broad including both public and private entities. The firm provides a variety of pre-design, design, and post-design services over a broad range of building typologies. For the purpose of this study the firm graciously supplied project data for approximately 400 projects completed over a period of several years. As described below several sets of similar projects were grouped in separate data sets and subjected to statistical analysis using each of the candidate production forms. Following a discussion of the results of the analysis the conclusions of this study are presented.

[258] Gordon Tullock, "A Plea to Economists Who Favor Liberty," *Eastern Economic Journal* 27, no. 2 (2001): 205. Tullock quoting Coase.

Data

The firm provided data for approximately 400 projects covering a multi-year period. The data consisted of man-power data by labor category by project plus the dollar value of the contract of each project. Organizing the data and creating the resulting data sets proceeded in five steps. The master data base, as provided by the firm, contained erroneous entries, errors in coding, and missing data; a condition that had to be corrected before proceeding further. A change in accounting systems, software upgrades, and deletion of certain records contributed to this situation. Additionally, some non-design activities were coded as design projects which tainted the file. The first step then became an exercise in identifying project records with missing or incomplete data elements and identification of non-design projects. Those projects deemed to fall into this category were purged from the data base.

A second scrub of the data base revealed inconsistent coding identifying clients, building typology, or service provided. Subsequently each project was examined to determine if these elements were correctly and consistently coded properly then corrections were applied to the data base as necessary.

The third scrub of the data base was made to parse the records into separate data sets based upon the coding of client, typology, and service provided. An evaluation of these data sets revealed that several lacked the sufficient number of observations needed to permit valid statistical analysis. These data sets were removed from further consideration.

An examination of the remaining projects revealed minor inconsistencies in coding labor categories. For example, an employee at a given level of experience appeared in more than one labor category. The category titles and the definitions of who fell into each one morphed over time as the data base evolved. Correcting these inconsistencies required a fourth scrub of the data base. At this point the final data base was constructed by grouping similar labor groups into representative labor categories. Category one consisted of junior architects while categories two and three were comprised of mid-level and senior architects respectively. The final category consists of employees performing various support staff functions. In addition to the manpower data the contractual value of the project was added to the data base to represent the output. The final data base consisted of five separate project/client/service data sets. As the statistical analysis began it soon was determined that two of data sets lacked the necessary number of observations and they also were removed from further consideration. The three remaining data sets are simply identified as data sets, one, two and three (n=20, 14, and 14 respectively).[259]

Analysis Methodology

Determining which production function(s) one, provides the best representation of the technologies employed by the firm and two, provides the soundest statistical basis for empirical study requires a two-pronged approach. The first approach employs strictly a statistical analysis. The second approach acknowledges that production functions are conceptual constructs representing the nature of the underlying technology and attempts to evaluate the statistical results in that light. In this manner both considerations of representation of technology and statistical viability are mutually considered in the final conclusion. The statistical approach is considered first.

The objective of the statistical analysis is identification of that production function model which produces the best goodness of fit with each data set. The procedure is to estimate the parameters of the five candidate production functions using each of the three project data sets

[259] This convention is adopted in accordance with the informed consent agreement signed between the firm and the research team to protect its identity.

and produce various measures of goodness of fit. A total of 12 sets of calculations (4 functions * three data sets) were performed utilizing the non-linear multiple regression function of IBM's SPSS® statistical software package version 20. SPSS employs two non-linear regression routines. The sequential quadratic programming (SQP) method was employed for each run as it provides non-linear optimization solutions of twice continuously differentiable equations with constraints.[260] The non-linear option in SPSS was chosen for two reasons. First, by using the same analytical process for each set of calculations a more valid and consistent comparison of similar measures of goodness of fit among all data sets is obtainable. Second, the SQP method of non-linear regression allows constraints to be placed on the parameters consistent with the underlying constraints of the production function, (e.g. some parameters are restricted to positive values). Normally ordinary least squares (OLS) linear regression estimations are sufficient for the Cobb-Douglas and Leontief production functions unless such analysis reveals coefficients or parameters that are less than zero. In those circumstances it is necessary to constrain coefficients and parameters to be equal to or greater than zero as negative values have no inherent economic value or meaning. Use of SQP, with parameters restricted to positive values, ensures that the results obtained retain an economic meaning. OLS cannot be utilized to estimate non-linear regression models and it was necessary to employ the SQP methodology with constraints for the Uzawa, Mukerji and Nerlove production functions. For consistency and to accommodate the need to constrain all coefficients and parameters the non-linear regression employing SQP was employed for these 12 runs. The exception to this process occurred with the Leontief production function. Lacking *a priori* information detailing the values of the coefficients of the various inputs in was necessary to estimate them. SPSS was used to conduct a series of linear regressions for each coefficient in all three data bases. Equipped with the values of the

[260] Paul T. Boggs and Jon W. Tolle, "Sequential Quadratic Programming," *Acta Numerica* 4(1995).

coefficients calculations were performed to yield the minimum output for each observation in each data set. The resulting minimum outputs were then treated as predicted outputs and were regressed against the actual output to determine the appropriate R^2.

The criteria for selection of the most statistically valid production function is a combination of the standard R^2 goodness of fit and the size of the standard error of each parameter. The principle goodness of fit criteria employed is the coefficient of determination, R^2. The nature of non-linear multiple regression analysis is such that R^2 is not a consistently reliable criterion as the value of R2 is often undefined. When R^2 is undefined the best measure of goodness of fit between two models becomes a comparison of the estimated value of individual parameters and their respective standard error.

The procedure outlined above whereby competing models are compared departs from standard model building procedures. It is common in most empirical research to test individual coefficients and parameters for significance using a standard t-test. In standard model building various candidate predictors are incorporated into a regression equation. The regression equation is then estimated using OLS or similar regression procedure. Independent variables that fail the standard t-test for the level of significance desired are eliminated. This process is continued until the highest R^2 is obtained. The revised model is then deemed the best model of behavior under study. For production functions the independent variables are the inputs observed in the production process. Removing variables (inputs) that fail a standard t-test from the production function effectively annuls the original production function and denies the validity of the original conceptualization of the underlying technology. As noted earlier in chapter three inputs are considered essential to the production process and cannot be removed without violating a basic premise of production economics. Here we resort to an examination of R^2 and when that proves insufficient the default becomes a comparison of the standard error of parameters.

The second approach is a conceptualization check. Each of the five candidate production functions describe certain economic characteristics of the underlying technology such as returns to scale for example. In the conceptualization check the question 'What is lost in describing the underlying technology by elimination of each candidate production function?'

The final judgment about which production function best represents the technologies employed by the firm and provides the soundest statistical basis for empirical study of other firms is a subjective judgment taking both approaches into account.

Analysis

Consistent with the procedures outlined above we begin with a statistical estimation and analysis of the candidate production functions. Here is presented an analysis of the 15 computer runs comparing and contrasting the estimates of coefficients, parameters, and measures of goodness of fit. The data is presented in the following five charts. Each chart depicts the results of the regression of one of the candidate production functions against each of the data sets.

Cobb-Douglas Production Function

The Cobb-Douglas production function is multiplicative in nature reflecting that an imperfect substitution condition exists. The $x_{i's}$ represent one of the labor categories discussed above. The A parameter is the scale parameter which describes the overall efficiency of the technology, the higher the value the more efficient the operation. The α exponents are the partial elasticities of each input, when summed they yield the return to scale of the technology. As shown in figure 15 below, The Cobb-Douglas model produces excellent R^2 values for data sets 1 and 2. The R^2 for data set 2 is undefined but the standard errors (numbers in parentheses) lay between those of data sets 1 and 2 except for the value of α_1. The value of the summation of the exponents, the return to scale, of 2.287, 1.729, and 1.46 for data sets 1-3 respectively indicate that the underlying technology for these building types reflect increasing returns to scale.

Table 2. Results of Regression Using the Cobb-Douglas Production Function

$y = A\prod_{1}^{n} x_i^{\alpha_i} = Ax_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}x_4^{\alpha_4}$			
Parameter/Data Set	1	2	3
R^2	.703	undef	.746
A	2.662 (37.841)	2.576 (19.52)	1.285 (3.895)
α_1	.690 (1.0)	1.425 (1.362)	.900 (.523)
α_2	.131 (.959)	.182 (.192)	.347 (.233)
α_3	1.313 (1.021)	.057 (.330)	.148 (.212)
α_4	.153 (2.421)	.065 (.372)	.065 (.276)

Leontief Production Function

The Leontief production function is additive in nature reflecting the recipe nature of the underlying technology. The Leontief production function reflects a constant return to scale technology. The β coefficients prescribe the fixed manner in which the inputs must be combined. As shown in figure 16 below, the R^2 for the first two data sets is moderately high while the R^2 for data set three is very high indicating an overall moderate to strong goodness of fit.

Table 3. Results of Regression Using the Leontief Production Function

$y=\min\left(\beta_1 x_1, \beta_2 x_2, \cdots, \beta_n x_n\right)=\beta_1 x_1, \beta_2 x_2, \beta_3 x_3, \beta_4 x_4$			
Parameter/Data Set	1	2	3
R^2	.667	.652	.949
β_1	306.372 (129.810)	120.913 (21.961)	119.395 (8.996)
β_2	365.275 (90.956)	223.460 (42.040)	1255.575 (260.949)
β_3	975.385 (106.529)	123.207 (48.301)	757.481 (148.796)
β_4	1731.027 (301.538)	629.329 (163.559)	457.869 (73.897)

Uzawa-McFadden CES Production Function

The Uzawa-McFadden CES production function is a complex representation of the underlying technology. At the heart of the function the β's are the coefficients of the inputs represented by the x's while the α exponent is the elasticity of substitution which in this case is fixed and constant. The A parameter is once again the efficiency scale factor and the v parameter yields the return to scale. As shown in figure 17 below, the R2 for data sets 1 and 2 are very high while undefined for data set 2. The standard errors for most of the parameters are extremely high for all three data sets indicative that the data lacks coherence. The value of v, the returns to scale indicator, is problematic as the standard error is many times greater than the value of the parameter. For this type of complex formulation one must be concerned regarding the number of observation required to adequately estimate all the parameters.

Table 4. Results of Regression Using the Uzawa-McFadden Function

$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha}\right)^{-\frac{v}{\alpha}} = A\left(\beta_1 x_1^{-\alpha} + \beta_2 x_2^{-\alpha} + \beta_3 x_3^{-\alpha} + \beta_4 x_4^{-\alpha}\right)^{-v/\alpha}$			
Parameter/Data Set	1	2	3
R^2	.823	undef	.885
A	1.125 (14412548.87)	1.245 (30600509.54)	8.074 (34606586.15)
β_1	.062 (63251)	.039 (151531.396)	.425 (4179.136)
β_2	.062 (63251.026)	.045 (173006.55)	.200 (1972.218)
β_3	.329 (336914.577)	.203 (786040.111)	.102 (1003.423)
β_4	.134 (137292.904)	.332 (1282382.074)	.267 (2627.596)
v	.091 (1.048)	.261 (8.426)	.003 (.969)
α	1.143 (.814)	1.66 (4.333)	1.280 (.296)

Mukerji Production Function

The Mukerji production function is the culmination of the development in CES functions. Here the modification is to permit unequal partial elasticities of substitution represented by the $\alpha_{i's}$. Interpretation is otherwise the same as for Uzawa-McFadden form except for the returns to scale determination. When $\sum_{i=1}^{n} \alpha_i = \alpha_0$ the return to scale is constant, when it is greater than α_0 the returns are increasing, and when less than α_0 returns are decreasing. As shown in figure 18, the Mukerji production exhibits widely varying R^2's. Additionally the estimates of all parameters vary greatly and their respective standard errors are extremely high. In this formulation with only four inputs the number of parameters to be estimated is very high at 10. Any increase in the number of inputs makes estimating this model very problematic.

Table 5. Results of Regression Using the Mukerji Function

$$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha_i}\right)^{-\frac{1}{\alpha_0}} = A\left(\beta_1 x_1^{\alpha_1} + \beta_2 x_2^{\alpha_2} + \beta_3 x_3^{\alpha_3} + \beta_4 x_4^{\alpha_4}\right)^{-1/\alpha_0}$$

Parameter/Data Set	1	2	3
R^2	.061	undef	.793
A	1.138 (380715839.0)	10.625 (391489228.9)	1.134 (9637308.562)
β_1	.258 (30806460.24)	4.236 (195919144.6)	.341 (331900.015)
α_1	.926 (67.279)	3.898 (65065.232)	.631 (13.896)
β_2	.250 (29841706.24)	.154 (7095963.571)	.228 (221937.328)
α_2	.907 (68.009)	2.953 (91.336)	.091 (1.942)
β_3	.283 (33792491.95)	.154 (7088274.76)	.088 (85326.915)
α_3	.855 (37.376)	2.289 (113.465)	.243 (2.218)
β_4	.252 (30075771.56)	.154 (7114888.643)	.307 (298255.823)
α_4	.911 (338.087)	2.189 (81.614)	.296 (9.586)
α_0	.357 (24.768)	1.252 (43.192)	.114 (2.313)

Nerlove-Ringstad Production Function

The Nerlove-Ringstad production function bears a striking resemblance to the Cobb-Douglas production function with the salient difference appearing in the elasticity of scale factor $1/(\gamma + \beta \ln Y)$. Evaluation of this term yields the returns to scale factor. Except for the return to scale factor interpretation of the Nerlove-Ringstad form is the same as for the Cobb-Douglas. As shown in figure 19 below, the R^2 values are relatively low for data sets 1 and 3 and undefined for data set 2. All Standard errors are extremely large. Given the complexity of this form its estimation and interpretation is problematic.

Table 6. Results of Regression Using the Nerlove-Ringstad Function

$y=\left(A\prod_{i=1}^{n} x_i^{\alpha_i}\right)^{\frac{1}{\gamma+\beta\ln y}} = \left(Ax_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}x_4^{\alpha_4}\right)^{1/(\gamma+\beta\ln y)}$			
Parameter/Data Set	1	2	3
R^2	.048	Undef	.535
A	2.645 (6622337.176)	2.149 (28.636)	3.166 (5182041.475)
α_1	9.419 (24244758.93)	7.476 (23431062.09)	42.662 (60613441.44)
α_2	4.357 (11214063.35)	7.004 (21951673.2)	26.179 (37194444.66)
α_3	7.945 (20452305.55)	7.076 (22177547.41)	26.485 (37628706.97)
α_4	3.769 (9701143.302)	2.025 (6345645.1)	33.755 (47958718.57)
γ	8.377 (21563461.75)	1.579 (4949190.096)	4.622 (6566458.981)
β	.173 (444058.918)	.951 (2980779.46)	4.665 (6627922.026)

In summary, the three data sets and five production functions analyzed here vary greatly in the degree to which they permit the construction of precise estimates of variables and reasonably tight confidence levels. In four of the five regressions data set two proved problematic returning undefined R^2 values and medium to very large standard errors. Only in case of the Leontief production function does data set two yield a defined R^2, .652, and exhibits moderate standard errors. Data sets one and three yield positive values for R^2 in each of the five regression scenarios but rather extreme standard errors in the cases of the Uzawa-McFadden, Mukerji, and Nerlove-Ringstad production functions. The most useful results obtain across all three data sets were for the regressions run under the Cobb-Douglas and Leontief production models.

The second requirement of the analysis is an assessment of which production function provides the best representation of the underlying technology. In this effort, advantage is taken of the previous statistical analysis. The Cobb-Douglas and Leontief production functions were clearly superior choices based upon the statistical analysis. Elimination of the Uzawa-

McFadden, Mukerji, and Nerlove-Ringstad production function is justified by the difficulties in their estimation and large standard errors. The difficulties in their estimation and interpretation make them unreliable models of the underlying technology of architectural firms. Retention of both the Cobb-Douglas and Leontief models is proposed as they represent two very different models of the underlying technology and thus provide the means of evaluating competing conceptualizations of the underlying technology of architectural firms. Their relative ease in estimation and interpretation are further justification.

Conclusion

Given the vicissitudes of the data used in this study and similar characteristics likely to be found in other data sets obtainable from functioning architectural firms the production functions best suitable for empirical analysis of architectural firms are the Cobb-Douglas and Leontief production functions. The differing manner in which these two productions treat the substitution of inputs in the production process make them interesting compliments to one another and should prove powerful analytical tools for architectural firms. Special note should be made of the difficulties in preparing a data base for statistical analysis. Care must be taken in construction of the data base to ensure consistency in coding and completeness of the data. When constructing final data sets for analysis all efforts must be made to ensure that only projects of similar characteristics are included in the data set. In the current effort several initial data sets had to be divided based upon the differing characteristics of projects. This serves to reduce the number of observations in each data set but results in model estimations with smaller errors.

CHAPTER 7

SUMMARY, CONCLUSION AND EPILOGUE

Summary

Concern over operating costs predates the first comprehensive study of their nature commission by AIA in 1966. Forty five years later the industry has yet to come fully to terms with the necessity of understanding the true nature of operating costs or to develop adequate analytical tools to determine their behavior. This dissertation introduced and developed the concept of 'cost structure' as it is applicable to architectural firms. The components of that cost structure are the physical costs in dollars, $time^2$ (manpower totals and time to completion), and scarce resources required to operate the firm but it also includes an understanding of cost behavior, the why and how costs vary according to building typology, services provided, client served, or level of design activity. Industry recognition of the requirement of understanding operating costs and adoption of analytical methods useful in understanding the firm's cost structure are vital to improving the financial picture of architectural firms. Hopefully this dissertation serves as an opening salvo in an ongoing conversation leading to permanent changes in how architects view this important aspect of the business side of architectural practice.

In this dissertation three analytical methods were introduced each with the potential to inform architects about the true nature the real costs of operating a firm – cost accounting, statistical analysis, and production economics. Two methods of analyzing costs, the cost accounting method, with emphasis on activity based accounting (ABC), and the statistical analysis method, with its emphasis on the creation of statistical measures of resource consumption, were briefly described. The main emphasis, however, remained focused on the use of production economics as the basis for an economic analysis of costs.

Production economics is about choice among alternatives, particularly choice that optimizes entrepreneur behavior be that behavior cost minimization or output maximization. The underlying premise is that the entrepreneur has discovered or developed the most efficient production technology possible. In this context technology represents the manner in which the productive inputs are brought together. Therefore the goal of production economics is about resource allocation, that is what inputs in what quantities minimizes the cost of operations for a given level of output, or maximizes output for a given cost or budget level. The production function, a mathematical representation of the arrangements of inputs necessary to achieve an output, is the vehicle by which optimization is achieved. The goal in this dissertation was to identify those production functions that best represented the technologies employed by an architectural firm while providing the soundest statistical basis for empirical study of other firms.

This dissertation enumerated and characterized production functions in a survey that spanned the period from 1928 and the introduction of the Cobb-Douglas production function to the end of the 20th century. In total 53 production functions and their variations were surveyed. Many of the forms were developed for macroeconomic applications and thus would prove unsuitable for our purposes. A number of the remaining forms were heavily parameterized and would prove unwieldy to estimate and difficult to interpret. Ultimately five production functions survived the winnowing process and were used in a case study of a mid-west architectural firm. Each of five production forms was estimated using three data sets supplied by the firm. Thus a total of fifteen computer runs utilizing IBM's SPSS® software program were compiled. The statistical results were then compared to determine if one or more production functions performed the best in representing the underlying technology with proving reasonably easy to estimate. Two production functions prove superior to the remaining three. Those production functions are the Cobb-Douglas and Leontief production functions.

The Cobb-Douglas production function represents an underlying technology in which tradeoffs are made between inputs to achieve optimal efficiency and minimum costs. The standard form of the Cobb-Douglas production is $y = A\prod_{1}^{n} x_i^{\alpha_i}$. The scale factor A is a scalar measure of the efficiency of the technology useful when comparing similar but different technologies available to the entrepreneur. The returns to scale factor can be calculated by summing the exponents associated with the inputs which when equal to one indicates a constant return to scale, when less than one it indicates decreasing returns to scale, and when greater than one it indicates increasing returns to scale. The basic information contained in the Cobb-Douglas formulation supports cost minimization or output maximization optimization routines. Cost minimization calculations support project budgeting and the bidding. Output maximization calculations support annual projections of resource requirements and annual output forecasts.

The Leontief production function differs significantly from the Cobb-Douglas form. The standard form of the Leontief production function is $y = \min(\beta_1 x_1, \beta_2 x_2)$. The Leontief presupposes that inputs are employed in a fixed ratio in the production of the final product. The goal of statistical analysis is simplified to the discovery or conformation of that ratio (although that ratio may be different for each building typology, client, or service provided). Determination of resources required for a given project or for annual production goals or annual resource requirement reduce to a simple linear extrapolation of that ratio. Cost minimization and out maximization have no relevance for technologies of the Leontief type.

Conclusion

The primary goal of this dissertation was the identification of those production functions that best represented the technologies employed by an architectural firm and provided the soundest statistical basis for empirical study of other firms. By the identification of the Cobb-Douglas and Leontief production function as those formulations suitable to these twin purposes the goal of this dissertation was achieved.

In the course of this investigation two unexpected conditions/issues were observed that warrant comment. The examination of the data base of project data for the medium sized mid-west based architectural firm used in this dissertation revealed interesting insight into the operations of architectural firms of that size. The first observation regards the impact on statistical analysis given the diversity of their portfolio. The firm undertook to design building, design modifications to existing building, and provide a diverse range of pre-design and post-design services across a broad range of building typologies and clients (commercial, public, and private organizations and individuals). The result of this diverse design portfolio is that some difficulty arose in classifying projects into discrete homogeneous data sets with the sufficiently large number of observations needed to support statistical analysis. This alone suggests that the production functions selected should yield to the simplest statistical methods while retaining the power to fully describe the underlying technology of the firm. As the portfolio of many medium sized firms is likely to exhibit a similar diversity then it is fair to assume that similar issues as found in our study firm exists also in the industry as well. The consequence may be that some firms, particularly small firms, simply may not have enough projects of a coherent homogeneous character to support economic analysis as detailed in this dissertation.

The second observation is more troubling yet promises a ray of good fortune. The standard error encountered in this study, particularly in data set 2 suggest that something anomalous may exist in the underlying technology employed in the projects comprising data set 2. One of the most basic premises of production economics is that the entrepreneur has developed and employs the most efficient production process available to the firm. Production economics then seeks to optimize the allocation of resources to achieve cost minimization or output maximization conditions. This relies upon the entrepreneur to consistently employ an efficient technological scheme. The large standard errors in data set 2 suggest that the underlying technology is not fully developed, not fully understood, or not consistently applied. One unanticipated outcome of an economic analysis of a given firm may well be to identify when these types of anomalous conditions exist and aid the entrepreneur in the discovery of a more efficient technology and point the way toward achieving a more profitable financial position.

Epilogue

We are left now to answer the question ‘Where do we go now?’ Two broad avenues avail themselves. The first leads to a pursuit of production economics issues. The second leads to a pursuit of the alternative forms of analysis necessary to a full understanding of the cost structure of the architectural firm.

In the pursuit of further research into production economics applicable to architectural firms several paths present themselves. The first path continues the basic research only just begun by this dissertation. This research should be extended to more firms of varying sizes, varying typologies, disparate client types, and possibly very different technologies.

A second path leads to the testing of a previous held assumption of production economics. That often unvoiced premise of production economics is that any production process can be represented and analyzed through its precepts. Key is the existence of an underlying technology. But what if there is no underlying coherent technology? What if whatever needs doing is done by whoever is available? How does the technology change both from small to large firms but also from small to large projects? Can the technology be captured by some analytical method and be fully described? Research into the applicability of production economics to firms of along the entire continuum from single person firms to large design bureaus, as advocated above, would be necessary to resolve this question.

A third area of potential research involves those aspects intentional omitted from this dissertation. Here the focus was upon the *n*-input single-output model when clearly the *m*-input multiple-output model presents an alternative analysis methodology. This may prove a very interesting line of investigation for architectural firms. Similarly, the application of frontier or data envelope analysis could prove very useful. In these analyses the basic assumption that the underlying technology is the most efficient is relaxed in favor of the discovery of more efficient technology schemes. Finally, application of duality theory may yield interesting insight into different forms of production functions applicable to architectural firms. In duality theory the cost function is seen as an alternative view of production possibilities. By reverse engineering the cost function the production function can be revealed. It would be very interesting to see if this has application in architecture.

The second broad avenue of additional research is found in the potential application of the alternative methods previously mentioned, namely the accounting and statistical methods of analysis. The accounting methodology may prove very insightful in the area of allocation of overhead costs to production activities, products, and clients. Similarly, the statistical analysis method holds great promise to provide insight into how and why changes in the characteristics of a building drive costs. The cost to design a large building is not necessarily twice or three times the cost to design a small building just because they are two or three times taller, wider, or have two or three time the square footage. A research project designed to investigate these potential goldmines of information may prove extremely profitable.

APPENDIX A

GLOSSARY OF ABBREVIATIONS

ABC	Activity based costing
AIA	American Institute of Architects
GAAP	Generally accepted accounting procedures
MRTS	Marginal rate of technical substitution
ROI	Return on investment

APPENDIX B

MATHEMATICAL NOTATIONS

y: Total output

x_i: Input factor i, i=1 to n

x_j: Input factor j, j=1 to n

$y = f(x) = f(x_1, x_2, ..., x_n)$: Long run production function, i.e. all inputs are treated as variable

$y = f(x) = (x_1 | x_2, x_3, ...x_n)$: Short run production functions with x_1 as the variable input

$f' = \frac{df}{dx}$: First derivative of a single variable function

$f_i = \frac{\delta y}{\delta x_i} = \frac{\delta f(x)}{\delta x_i}$: First partial derivative of the function $f(x)$ with respect to input x_i

$f_{ij} = \frac{\delta^2 y}{\delta x_i \delta x_j} = \frac{\delta f^2(x)}{\delta x_i \delta x_j}$: Second partial derivative of the function $f(x)$ with respect to input x_i and x_j

dx_i = The change in value of the x_i variable

A = A scalar leading coefficient (example $y = Ax_1^{\alpha_1} x_2^{\alpha_2}$)

α = exponent of an input (example $y = Ax_1^{\alpha_1} x_2^{\alpha_2}$)

β = coefficient of an input (example $y = \beta_1 x_1 + \beta_2 x_2$)

σ = elasticity of substitution

e = natural logarithm

$\delta, \gamma, \rho, \upsilon$ = additional parameters

APPENDIX C

CONCAVITY AND CONVEXITY

Introduction

This appendix details the procedures for the tests of concavity and convexity for both the 2-input case and the *n*-input case. The test of convexity for the 2-input case is useful in determining the convexity of isoquants, whereas the more general *n*-input test for global concavity is useful in determine whether the underlying production function meets the criteria of concavity required of production functions.

Concavity and convexity: the 2-input case[261]

For a twice differentiable production function with continuous partial derivatives of first and second order, concavity and convexity are determine by the sign of the second partial derivative of each variable input and the sign of the Hessian matrix of second partial derivatives as follows:

a.) $f(x)$ is concave provided that $f_{11} \le 0,\ f_{22} \le 0$ and $\left[\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}\right] \ge 0$

b.) $f(x)$ is convex provided that $f_{11} \ge 0,\ f_{22} \ge 0$ and $\left[\begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}\right] \ge 0$

[261] Knut Sydsaeter and Peter J. Hammond, *Mathematics for Economic Analysis* (Englewood Cliffs, New Jersey: Prentice Hall, 1995). 632.

Concavity and convexity: the n-input case[262]

Given a twice differentiable continuous function of the form $y = f(x_1, \ldots, x_n)$ the concavity or convexity of the function can be determined by examination of the principal minors of the Hessian matrix of second partial derivatives. The Hessian matrix is formed of all the partial second derivatives of the function $f(x)$ as follows:

$$H(f(x)) = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{pmatrix} \text{ where } f_{ij} = \frac{\delta^2 f(x)}{\delta x_i \delta x_j},\ i = 1, \ldots, n,\ j = 1, \ldots, n$$

The *n* principle minors of the Hessian matrix are:

$$|H_1| = |f_{11}|\ ,\ |H_2| = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix},\ |H_3| = \begin{vmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{vmatrix}, \ldots,\ |H_n| = \begin{vmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{n1} & f_{n1} & \cdots & f_{nn} \end{vmatrix}$$

The function $y = f(x_1, \ldots, x_n)$ is said to be concave everywhere if $|H_1| \leq 0$, $|H_2| \geq 0$, $|H_3| \leq 0$, $\ldots$, $(-1)^n |H_n| \geq 0$, and is said to convex everywhere if $|H_1| \geq 0$, $|H_2| \geq 0$, $|H_3| \geq 0$, $\ldots$, $|H_n| \geq 0$

[262] Ibid., 637-38.

APPENDIX D

ELASTICITY OF SUBSTITUTION

Introduction

This appendix derives the formula for the elasticity of substitution and demonstrates the calculation of the elasticity of substitution for the Cobb-Douglas production function in the two input case where $y = Ax_1^{\alpha_1} x_2^{\alpha_2}$.

Derivation of the formula for the Elasticity of Substitution[263]

The elasticity of substitution σ is defined as the proportionate rate of change of the input ratio divided by the proportionate rate of change in the MRTS,

$$\sigma = \left(\frac{d\left(\frac{x_2}{x_1}\right)}{\frac{x_2}{x_1}} \right) \Bigg/ \left(\frac{d\left(\frac{f_1}{f_2}\right)}{\left(\frac{f_1}{f_2}\right)} \right) \tag{D.1}$$

By inverting the numerator and multiplying through we obtain,

$$\sigma = \left(\frac{\frac{f_1}{f_2}}{\frac{x_2}{x_1}} \right) * \left(\frac{d\left(\frac{x_2}{x_1}\right)}{d\left(\frac{f_1}{f_2}\right)} \right) \tag{D.2}$$

Proceed by first evaluating the total differential $d\left(\frac{x_2}{x_1}\right)$,

$$\frac{\partial\left(\frac{x_2}{x_1}\right)}{\partial x_1} dx_1 + \frac{\partial\left(\frac{x_2}{x_1}\right)}{\partial x_2} dx_2 \tag{D.3}$$

[263] Nobel Media AB, 29-30. The derivation here is the same as appearing on pages 29-30.

Next take the derivative,

$$\frac{-x_2}{x_1^2}dx_1 + \frac{-x_1}{x_1^2}dx_2 \tag{D.4}$$

By collecting terms obtain,

$$\frac{-x_2 dx_1 + x_1 dx_2}{x_1^2}. \tag{D.5}$$

From the formula for the MRTS we know that $dx_2 = -\left(\frac{f_1}{f_2}\right)dx_1$ and by substituting that into equation (E.5) obtain,

$$\frac{-x_2 dx_1 + x_1\left(\frac{-f_1}{f_2}dx_1\right)}{x_1^2}. \tag{D.6}$$

By collecting terms obtain,

$$\frac{\left(-x_2 - \frac{x_1 f_1}{f_2}\right)dx_1}{x_1^2}. \tag{D.7}$$

Next consider the total differential $d\left(\frac{f_1}{f_2}\right)$,

$$\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_1}dx_1 + \frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_2}dx_2. \tag{D.8}$$

From the formula for the MRTS we know that $dx_2 = -\left(\frac{f_1}{f_2}\right)dx_1$ and by substituting that into equation(D.8) obtain,

$$\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_1}dx_1 + \frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_2} * \frac{-f_1}{f_2}dx_1. \tag{D.9}$$

By collecting terms obtain,

$$\left[\frac{\partial \frac{f_1}{f_2}}{\partial x_1} - \left(\frac{f_1}{f_2}\right)\frac{\partial \frac{f_1}{f_2}}{\partial x_2}\right]dx_1. \tag{D.10}$$

By substituting equations (D.7) and (D.10) into equation(D.2) obtain,

$$\sigma = \left(\frac{\frac{f_1}{f_2}}{\frac{x_2}{x_1}}\right) * \left(\frac{\frac{\left(-x_2 - \frac{x_1 f_1}{f_2}\right)dx_1}{x_1^2}}{\left[\frac{\partial \frac{f_1}{f_2}}{\partial x_1} - \left(\frac{f_1}{f_2}\right)\frac{\partial \frac{f_1}{f_2}}{\partial x_2}\right]dx_1}\right). \tag{D.11}$$

By collecting and rearranging terms obtain first,

$$\frac{\frac{f_1}{f_2}}{x_1^2\left(\frac{x_2}{x_1}\right)} * \frac{-\left(x_2 + x_1\left(\frac{f_1}{f_2}\right)\right)}{-\left[\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_2} * \frac{f_1}{f_2} - \frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_1}\right]} \tag{D.12}$$

then,

$$\frac{f_1\left(f_1x_1+f_2x_2\right)}{f_2^2x_1x_2\left[\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_2}*\left(\frac{f_1}{f_2}\right)-\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_1}\right]} \tag{D.13}$$

and finally,

$$\frac{f_1\left(f_1x_1+f_2x_2\right)}{f_2x_1x_2\left[f_1\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_2}-f_2\frac{\partial\left(\frac{f_1}{f_2}\right)}{\partial x_1}\right]}. \tag{D.14}$$

By evaluating the term in brackets in the denominator obtain,

$$f_1\left(\frac{f_2f_{12}-f_1f_{22}}{f_2^2}\right)-f_2\left(\frac{f_2f_{11}-f_1f_{21}}{f_2^2}\right). \tag{D.15}$$

By expanding terms obtain,

$$\frac{f_1f_2f_{12}-f_1^2f_{22}-f_2^2f_{11}+f_1f_2f_{21}}{f_2^2}. \tag{D.16}$$

By collecting terms obtain,

$$\frac{2f_1f_2f_{12}-f_1^2f_{22}-f_2^2f_{11}}{f_2^2}. \tag{D.17}$$

By substituting equation(D.17) into the brackets in equation(D.14) and collecting terms obtain the final expression for the elasticity of substitution as,

$$\sigma=\frac{f_1f_2\left(f_1x_1+f_2x_2\right)}{x_1x_2\left(2f_1f_2f_{12}-f_1^2f_{22}-f_2^2f_{11}\right)}. \tag{D.18}$$

Elasticity of Substitution: Cobb-Douglas Production Function

The formula for the elasticity of substitution as derived above is,

$$\sigma = \frac{f_1 f_2 \left(f_1 x_1 + f_2 x_2 \right)}{x_1 x_2 \left(2 f_{12} f_1 f_2 - f_1^2 f_{22} - f_2^2 f_{11} \right)}.$$ [264]

Begin by taking the first, second, and cross-partial derivatives for the Cobb-Douglas production function given by,

$$y = Ax_1^{\alpha_1} x_2^{\alpha_2} \tag{D.19}$$

$$f_1 = \alpha_1 A x_1^{\alpha_1 - 1} x_2^{\alpha_2} \tag{D.20}$$

$$f_2 = \alpha_2 A x_1^{\alpha_1} x_2^{\alpha_2 - 1} \tag{D.21}$$

$$f_1^2 = \alpha_1^2 A^2 x_1^{2\alpha_1 - 2} x_2^{2\alpha_2} \tag{D.22}$$

$$f_2^2 = \alpha_2^2 A^2 x_1^{2\alpha_1} x_2^{2\alpha_2 - 2} \tag{D.23}$$

$$f_{11} = \alpha_1 \left(\alpha_1 - 1 \right) A x_1^{\alpha_1 - 2} x_2^{\alpha_2} \tag{D.24}$$

$$f_{12} = \alpha_1 \alpha_2 A x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1} \tag{D.25}$$

$$f_{22} = \alpha_2 \left(\alpha_2 - 1 \right) A x_1^{\alpha_1} x_2^{\alpha_2 - 2}. \tag{D.26}$$

Begin by evaluating the first two terms of the numerator $f_1 f_2$,

$$f_1 f_2 = \left(\alpha_1 A x_1^{\alpha_1 - 1} x_2^{\alpha_2} \right) \left(\alpha_2 A x_1^{\alpha_1} x_2^{\alpha_2 - 1} \right). \tag{D.27}$$

Then by multiplying through, collecting terms, and substituting *y* from equation (D.19) where appropriate obtain,

$$\alpha_1 \alpha_2 y^2 x_1^{-1} x_2^{-1}. \tag{D.28}$$

[264] Ibid., 30.

Evaluating the last term in the numerator $\left(f_1 x_1 + f_2 x_2\right)$ obtain,

$$\left(\alpha_1 A x_1^{\alpha_1 - 1} x_2^{\alpha_2}\right)\left(x_1\right) + \left(\alpha_2 A x_1^{\alpha_1} x_2^{\alpha_2 - 1}\right)\left(x_2\right). \tag{D.29}$$

By collecting terms and substituting y from equation (D.19) where appropriate obtain,

$$\alpha_1 y + \alpha_2 y. \tag{D.30}$$

Combining terms from both equation (D.28) and (D.30) obtain the numerator,

$$\alpha_1 \alpha_2 y^2 x_1^{-1} x_2^{-1}\left(\alpha_1 y + \alpha_2 y\right). \tag{D.31}$$

In evaluating the denominator start with the expression $\left(2 f_{12} f_1 f_2 - f_1^2 f_{22} - f_2^2 f_{11}\right)$ beginning with the first term,

$$2 f_{12} f_1 f_2 = (2)\left(\alpha_1 \alpha_2 A x_1^{\alpha_1 - 1} x_2^{\alpha_2 - 1}\right)\left(\alpha_1 A x_1^{\alpha_1 - 1} x_2^{\alpha_2}\right)\left(\alpha_2 A x_1^{\alpha_1} x_2^{\alpha_2 - 1}\right). \tag{D.32}$$

By collecting terms and substituting y from equation(D.19) as appropriate obtain,

$$2\alpha_1^2 \alpha_2^2 y^3 x_1^{-2} x_2^{-2}. \tag{D.33}$$

Next evaluate the second term,

$$f_1^2 f_{22} = \left(\alpha_1^2 A^2 x_1^{2\alpha_1 - 2} x_2^{2\alpha_2}\right)\left(\alpha_2\left(\alpha_2 - 1\right) A x_1^{\alpha_1} x_2^{\alpha_2 - 2}\right). \tag{D.34}$$

By collecting terms and substituting y from equation(D.19) as appropriate obtain,

$$\left(\alpha_1^2 \alpha_2^2 - \alpha_1^2 \alpha_2\right) y^3 x_1^{-2} x_2^{-2}. \tag{D.35}$$

Next evaluate the last term,

$$f_1^2 f_{22} = \left(\alpha_1^2 A^2 x_1^{2\alpha_1 - 2} x_2^{2\alpha_2}\right)\left(\alpha_2\left(\alpha_2 - 1\right) A x_1^{\alpha_1} x_2^{\alpha_2 - 2}\right). \tag{D.36}$$

And by collecting terms and substituting y from equation(D.19) as appropriate obtain,

$$\left(\alpha_1^2\alpha_2^2-\alpha_1^2\alpha_2\right)y^3x_1^{-2}x_2^{-2}. \tag{D.37}$$

The denominator is,

$$x_1x_2\left(2f_{12}f_1f_2-f_1^2f_{22}-f_2^2f_{11}\right). \tag{D.38}$$

By substituting equations (D.33), (D.35), and (D.37) obtain,

$$x_1x_2\left(\left[2\alpha_1^2\alpha_2^2y^3x_1^{-2}x_2^{-2}\right]-\left[\left(\alpha_1^2\alpha_2^2-\alpha_1^2\alpha_2\right)y^3x_1^{-2}x_2^{-2}\right]-\left[\left(\alpha_1^2\alpha_2^2-\alpha_1\alpha_2^2\right)y^3x_1^{-2}x_2^{-2}\right]\right). \tag{D.39}$$

By collecting terms and substituting y from equation(D.19) where appropriate obtain,

$$x_1^{-1}x_2^{-1}\left[\alpha_1\alpha_2y^3\right]\left[\alpha_1+\alpha_2\right]. \tag{D.40}$$

Combining the numerator equation (D.31) and denominator equation(D.40) obtain,

$$\frac{\alpha_1\alpha_2y^2x_1^{-1}x_2^{-1}\left(\alpha_1y+\alpha_2y\right)}{x_1^{-1}x_2^{-1}\left[\alpha_1\alpha_2y^3\right]\left[\alpha_1+\alpha_2\right]}. \tag{D.41}$$

By collecting terms obtain,

$$\frac{\alpha_1\alpha_2y^3x_1^{-1}x_2^{-1}\left(\alpha_1+\alpha_2\right)}{\alpha_1\alpha_2y^3x_1^{-1}x_2^{-1}\left[\alpha_1+\alpha_2\right]}=1. \tag{D.42}$$

The demonstration above shows that the elasticity of substitution for the Cobb-Douglas production formula is equal to one.

APPENDIX E

OPTIMIZATION EXAMPLES

Introduction

This appendix presents demonstrations of cost minimization under constrained output and output maximization under constrained cost utilizing the Lagrangian optimization technique. Both examples employ the same production and cost functions. For illustration purposes a hypothetical production function of $y = 10x_1^{.6}x_2^{.4}$ and a cost function of $C_0 = w_1x_1 + w_2x_2 + FC$, where $w_1 = \$10$, $w_2 = \$15$, and fixed costs of \$25 are employed.

Cost Minimization under Constrained Output Conditions

In this example the goal is minimizing cost given an output of 500 units. Utilizing the Lagrangian method of optimization we begin by formulating the Lagrangian function as,

$$L = w_1x_1 + w_2x_2 + FC + \lambda\left(500 - 10x_1^{.6}x_2^{.4}\right). \tag{E.1}$$

Begin by reformulating equation (E.1) by inserting the cost values and performing a logarithmic transformation of the production function obtaining,

$$L = 10x_1 + 15x_2 + 25 + \lambda\left(6.2146 - 2.3026 - .6\ln x_1 - .4\ln x_2\right). \tag{E.2}$$

Next, solve for the first order conditions yielding,

$$\frac{\partial L}{\partial x_1} = 10 - \frac{.6\lambda}{x_1} = 0, \tag{E.3}$$

$$\frac{\partial L}{\partial x_2} = 15 - \frac{.4\lambda}{x_2} = 0, \tag{E.4}$$

and $$\frac{\partial L}{\partial \lambda} = \left(6.2146 - 2.3026 - .6\ln x_1 - .4\ln x_2\right) = 0. \quad \text{(E.5)}$$

Solve equations (E.3) and (E.4) for λ yielding,

$$\lambda = 16.6667 x_1, \quad \text{(E.6)}$$

and $$\lambda = 37.5 x_2. \quad \text{(E.7)}$$

Use equations (E.6) and (E.7) to solve for x_1,

$$x_1 = 2.25 x_2. \quad \text{(E.8)}$$

Insert the value for x_1 into equation(E.5),

$$6.2146 - 2.3026 - .6\ln\left(2.25 x_2\right) - .4\ln x_2 = 0. \quad \text{(E.9)}$$

Solve equation (E.9) for x_2 by first collecting the x_2's on the left and the constants on the right to obtain,

$$-.6\ln\left(2.25 x_2\right) - .4\ln x_2 = -3.912. \quad \text{(E.10)}$$

Multiple both sides of equation (E.10) by -1 and perform an exponential transformation yielding,

$$\left(2.25 x_2\right)^{.6} x_2^{.4} = 49.9988. \quad \text{(E.11)}$$

Then in equation (E.11) expand the term $(2.25 x_2)^{.6}$, combine x_2 terms and divide both sides by the constant 1.6267, $\left(2.25^{.6} = 1.6267\right)$, to obtain,

$$x_2^1 = 30.7362. \quad \text{(E.12)}$$

Insert the value of x_2 back into equation (E.8) to determine x_1,

$$x_1 = 2.25 x_2 = 2.25 * 30.7362 = 69.1565 \quad \text{(E.13)}$$

The second order conditions require the determinant of the bordered Hessian matrix be greater than zero as follows,

$$H = \begin{bmatrix} -0.0064 & 0 & -10 \\ 0 & -0.01 & -15 \\ -10 & -15 & 0 \end{bmatrix} = 2.44 \tag{E.14}$$

Returning to the original production function and substituting the values for x_1 and x_2 calculated in equations (E.12) and (E.13) yielding,

$$y = 10(69.1565)^{.6}(30.7362)^{.4} = 500, \tag{E.15}$$

thus demonstrating the correctness of the values of x_1 and x_2. To obtain the minimum cost for producing 500 units substitute the values of x_1 and x_2 back into the cost function yielding,

$$C_0 = \$10(69.1565) + \$15(30.7362) + 25 = \$1{,}177.61. \tag{E.16}$$

Output Maximization under Constrained Cost Conditions

In this example the goal is maximizing output given a budget of \$1200. Utilizing the Lagrangian method of optimization we begin by formulating the Lagrangian function as,

$$L = 2.3026 + .6\ln x_1 + .4\ln x_2 + \lambda(1200 - 10x_1 - 15x_2 - 25). \tag{E.17}$$

Begin by solving for the first order conditions which yields,

$$\frac{\partial L}{\partial x_1} = \frac{.6}{x_1} - 10\lambda = 0, \tag{E.18}$$

$$\frac{\partial L}{\partial x_2} = \frac{.4}{x_2} - 15\lambda = 0, \tag{E.19}$$

and $$\frac{\partial L}{\partial \lambda} = 1200 - 10x_1 - 15x_2 - 25 = 0\,. \qquad \text{(E.20)}$$

Solving equation (E.18) and (E.19) for λ yields,

$$\lambda = .06x_1^{-1} \qquad \text{(E.21)}$$

$$\lambda = .0267x_2^{-1} \qquad \text{(E.22)}$$

Given that both equations (E.21) and (E.22) are set equal to λ the right side of each expression is equal to each other, restated that yields,

$$.06x_1^{-1} = .0267x_2^{-1}\,. \qquad \text{(E.23)}$$

By solving for x_1 we obtain,

$$x_1 = 2.2472x_2\,. \qquad \text{(E.24)}$$

Substituting this value for x_1 into equation (E.20) yields,

$$1200 - 10\left(2.2472x_2\right) - 15x_2 - 25 = 0 \qquad \text{(E.25)}$$

Solving for x_2 yields,

$$x_2 = 31.3567 \qquad \text{(E.26)}$$

Substituting the value of x_2 from equations (E.26) and solving equation (E.24) for x_1 yields,

$$x_1 = 2.2472x_2 = 2.2472(31.3567) = 70.4648 \qquad \text{(E.27)}$$

The second order conditions require the determinant of the bordered Hessian matrix be greater than zero as follows,

$$H = \begin{bmatrix} -0.00621 & 0 & -10 \\ 0 & -0.01 & -15 \\ -10 & -15 & 0 \end{bmatrix} = 2.39725\,. \qquad \text{(E.28)}$$

Solving equation (E.20) using the values of x_1 and x_2 from equations (E.27) and (E.26) respectively we verify our solution,

$$1200 - 10(70.4648) - 15(31.3567) - 25 = 0 \tag{E.29}$$

Substituting the values calculated for x_1 and x_2 into the production function we obtain the maximum output for a budget of \$1200,

$$y = Ax_1^{.6}x_2^{.4} = 10(70.4648)^{.6}(31.3567)^{.4} = 509.701 \text{ units} \tag{E.30}$$

APPENDIX F

DERIVATION OF THE ACMS CONSTANT ELASTICITY OF SUBSTITUION PRODUCTION FUNCTION

Introduction

The purpose of this appendix is twofold. The primary purpose is the presentation of a detailed step-by-step replication of the derivation of the ACMS Constant Elasticity of Substitution production function as it appears in *Capital-Labor Substitution and Economic Efficiency*[265]. The original article omits or combines a number of intermediate steps. The expanded derivation is an attempt to improve clarity and enhance the reader's comprehension of the development of the ACMS production function. The secondary purpose is to demonstrate the process whereby many theorist have begun with a regression model, converted it to a differential equation by some transformation and then proceeded to solve the differential equation to produce a new production function. While this process was first employed by Kenneth Arrow and his colleagues it was later employed by other economists with considerable success.

Development of the ACMS Production Function

Two widely held theories regarding the elasticity of substitution have already been mentioned in chapter four. The Cobb-Douglas production function in its original form presents the first case, where the exponents sum to one or $\sigma = 1$. The Leontief production function represents the second case where no substitution is permitted thus $\sigma = 0$. Arrow notes that in

[265] Arrow et al., "Capital-Labor Substitution and Economic Efficiency."

turning to empirical evidence it is apparent that varying degrees of substitution exist. The study leading to the development of the ACMS production function was an attempt to derive a mathematical function that reflected their observed production numbers and which had the following three properties: homogeneity, constant elasticity of substitution between capital and labor, and allowed for the possibility of varying elasticities between industries.

Arrow's group collected and analyzed data from 19 counties, the number of industries in each country varied from a low of 2 industries and ranged up to a maximum of 24 industries. They tested two competing regression equations,

$$\frac{V}{L} = c + dW + \eta \tag{F.1}$$

and

$$\log\frac{V}{L} = \log a + b\log W + \varepsilon . \tag{F.2}$$

Such that, V : value added in thousand of U.S. dollars

L : Labor input in man-years

and W : Money wage rate (total labor cost divided by L) in dollars per man-year.

They concluded that while both regression equations gave good fits to observed data, the logarithmic form (F.2) performed somewhat better. The derivation is as follows.

Given that the underlying production function may be written as,

$$V = F\left(K, L\right), \tag{F.3}$$

and assuming that the function is homogeneous of degree one, then the production function may be rewritten as,

$$\frac{V}{L} = F\left(\frac{K}{L}, 1\right). \tag{F.4}$$

By substituting y for $\frac{V}{L}$ and x for $\frac{K}{L}$ the production function may be rewritten again as,

$$y = f(x). \tag{F.5}$$

Correspondingly the marginal product of labor becomes,

$$f(x) - xf'(x). \tag{F.6}$$

Under profit maximizing conditions and perfect competition, the wage rate must equal the marginal product of labor,

$$w = f(x) - xf'(x) = y - x\frac{dy}{dx}. \tag{F.7}$$

From equation (F.1) we see that y and w enjoy a functional relationship,

$$y = g(x) \tag{F.8}$$

Substituting equations (F.5) and the right hand side of equation (F.7) into (F.8) results in,

$$y = g\left(y - x\frac{dy}{dx}\right). \tag{F.9}$$

Therefore by substituting the right hand side of equation (F.7) into equation (F.2) we obtain the differential equation,

$$\log\frac{V}{L} = \log a + b\log\left(y - x\frac{dy}{dx}\right). \tag{F.10}$$

Taking the antilogarithms of equation (F.10) we obtain,

$$y = a\left(y - x\frac{dy}{dx}\right)^{b}. \tag{F.11}$$

By algebraic manipulation first divide through by a and take the log of both sides to obtain,

$$\log\left(\frac{y}{a}\right) = b\log\left(y - x\frac{dy}{dx}\right), \tag{F.12}$$

then multiple both sides by $\frac{1}{b}$ and take the antilogarithms of both sides to obtain,

$$\frac{y^{1/b}}{a^{1/b}} = y - x\frac{dy}{dx}. \tag{F.13}$$

Subtracting y from both sides and multiplying both sides by -1 yields,

$$y - \frac{y^{1/b}}{a^{1/b}} = x\frac{dy}{dx}. \tag{F.14}$$

Expressing the left hand side of equation (F.14) as a common fraction and divide both sides by x yields,

$$\frac{a^{1/b}y - y^{1/b}}{a^{1/b}x} = \frac{dy}{dx}. \tag{F.15}$$

Factoring out $a^{1/b}$ in the numerator of the left hand side yields,

$$\frac{a^{1/b}\left(y - a^{-1/b}y^{1/b}\right)}{a^{1/b}x} = \frac{dy}{dx}. \tag{F.16}$$

Cancelling out the term $a^{1/b}$ in both the numerator and denominator yields,

$$\frac{y - a^{-1/b}y^{1/b}}{x} = \frac{dy}{dx}. \tag{F.17}$$

Factoring out y in the numerator yields,

$$\frac{y\left(1 - a^{-1/b}y^{1/b-1}\right)}{x} = \frac{dy}{dx}, \tag{F.18}$$

Substituting $\alpha = a^{1/b}$ and $\rho = \frac{1}{b} - 1$ into equation (F.18) yields,

$$\frac{y\left(1 - \alpha y^{\rho}\right)}{x} = \frac{dy}{dx}. \tag{F.19}$$

Multiplying equation (F.19) by dx and $1 \big/ y\left(1-\alpha y^{\rho}\right)$ yields,

$$\frac{dx}{x} = \frac{dy}{y\left(1-\alpha y^{\rho}\right)}. \tag{F.20}$$

In the next several steps equation (F.20) is transformed into equation (F.31) by a partial-fraction expansion of the right hand side of equation (F.20) as follows. Construct an equation using the right hand side of equation (F.20) and a separation of that expression into the sum of two fractions with unknown numerators A and B as follows,

$$\frac{dy}{y\left(1-\alpha y^{\rho}\right)} = \frac{A}{y} + \frac{B}{\left(1-\alpha y^{\rho}\right)}. \tag{F.21}$$

Multiple both sides of equation (F.21) by the common lowest denominator yielding,

$$dy = A\left(1-\alpha y^{\rho}\right) + By. \tag{F.22}$$

The general procedure at this point is to solve this equation in terms of both A and B thus constructing a new expression that is equivalent to the left hand side of equation(F.21). By inspection of equation (F.21) it is reasonable to conclude that either A and/or B must include the term dy. For convenience begin by setting $A = dy$ then making the appropriate substitution and expansion of the term containing A to obtaining,

$$dy = (dy) - (dy)\alpha y^{\rho} + By. \tag{F.23}$$

We begin by subtracting the $(dy) - (dy)\alpha y^{\rho}$ term from both sides and combining terms to obtain,

$$(dy)\alpha y^{\rho} = By. \tag{F.24}$$

Dividing both sides by y yields,

$$\frac{(dy)\alpha y^{\rho}}{y} = B. \tag{F.25}$$

The equivalent of equation (F.25) is,

$$(dy)\alpha y^{\rho-1} = B. \tag{F.26}$$

Substituting the value of B from equation (F.26) into (F.22) obtaining,

$$dy = A\left(1-\alpha y^{\rho}\right) + (dy)\alpha y^{\rho-1} y. \tag{F.27}$$

Begin to solve for A by combining terms and subtracting $(dy)\alpha y^{\rho}$ from both sides yielding,

$$dy - (dy)\alpha y^{\rho} = A\left(1-\alpha y^{\rho}\right). \tag{F.28}$$

Factor out a dy in the left hand side and divide both sides by $\left(1-\alpha y^{\rho}\right)$ obtaining,

$$\frac{dy\left(1-\alpha y^{\rho}\right)}{\left(1-\alpha y^{\rho}\right)} = A = dy. \tag{F.29}$$

Thus confirming our original value for A. Next substitute the values for A and B into equation (F.21) obtaining,

$$\frac{dy}{y\left(1-\alpha y^{\rho}\right)} = \frac{dy}{y} + \frac{\alpha y^{\rho-1} dy}{\left(1-\alpha y^{\rho}\right)} \tag{F.30}$$

The right hand side of (F.30) is the partial fraction decomposed equivalent of the left hand side.

We may now rewrite equation (F.20) as,

$$\frac{dx}{x} = \frac{dy}{y} + \frac{\alpha y^{\rho-1} dy}{\left(1-\alpha y^{\rho}\right)}. \tag{F.31}$$

Integration of equation (F.31) yields,

$$\log x = \log y - \frac{1}{\rho}\log\left(1-\alpha y^{\rho}\right) + \frac{1}{\rho}\log \beta . \tag{F.32}$$

By taking the antilogarithms we obtain,

$$x = \frac{y\beta^{1/\rho}}{\left(1-\alpha y^{\rho}\right)^{1/\rho}} . \tag{F.33}$$

Raising both sides by ρ yields,

$$x^{\rho} = \frac{y^{\rho}\beta}{\left(1-\alpha y^{\rho}\right)} . \tag{F.34}$$

In the next several steps we must first solve equation (F.34) for y^{ρ} then y in turn; we begin by isolating the y terms on the left hand side of equation (F.35) as follows,

$$\frac{y^{\rho}}{\left(1-\alpha y^{\rho}\right)} = \frac{x^{\rho}}{\beta} . \tag{F.35}$$

Factor y^{ρ} out of the denominator of the left hand side of equation (F.35) obtaining,

$$\frac{y^{\rho}}{y^{\rho}\left(y^{-\rho} - \alpha\right)} = \frac{x^{\rho}}{\beta} . \tag{F.36}$$

Cancel the y^{ρ} then invert the equation to obtain,

$$\left(y^{-\rho} - \alpha\right) = \frac{\beta}{x^{\rho}} . \tag{F.37}$$

Adding α to both sides yields,

$$y^{-\rho} = \frac{\beta}{x^{\rho}} + \alpha . \tag{F.38}$$

Adding the fractions on the right hand side yields,

$$y^{-\rho} = \frac{\beta + \alpha x^{\rho}}{x^{\rho}}. \tag{F.39}$$

Taking the log of both sides yields,

$$-\rho \log y = \log\left(\beta + \alpha x^{\rho}\right) - \rho \log x. \tag{F.40}$$

Multiplying both sides by $-\frac{1}{\rho}$ yields,

$$\log y = -\frac{1}{\rho}\log\left(\beta + \alpha x^{\rho}\right) + \log x. \tag{F.41}$$

Taking the antilogarithms of both sides yields,

$$y = \left(\beta + \alpha x^{\rho}\right)^{-\frac{1}{\rho}} x, \tag{F.42}$$

A more useful form of equation (F.42) is found as follows, by raising both sides by $-\rho$

obtaining,

$$y^{-\rho} = \left(\beta + \alpha x^{\rho}\right) x^{-\rho}. \tag{F.43}$$

Multiplying through on the right hand side obtain,

$$y^{-\rho} = \left(\beta x^{-\rho} + \alpha\right). \tag{F.44}$$

Raising both sides by $-\frac{1}{\rho}$ yields,

$$y = \left(\beta x^{-\rho} + \alpha\right)^{-\frac{1}{\rho}}. \tag{F.45}$$

To write out the full production function we substitute the original values of $y = \frac{V}{L}$ and $x = \frac{K}{L}$ to

obtain,

$$\frac{V}{L} = \left(\beta\left(\frac{K}{L}\right)^{-\rho} + \alpha\right)^{-\frac{1}{\rho}}. \tag{F.46}$$

Then multiple through by L to obtain,

$$V = L\left(\beta\left(\frac{K}{L}\right)^{-\rho} + \alpha\right)^{-1/\rho}. \tag{F.47}$$

Raising both sides by $-\rho$ yields,

$$V^{-\rho} = L^{-\rho}\left(\beta\left(\frac{K}{L}\right)^{-\rho} + \alpha\right). \tag{F.48}$$

Multiplying through the right hand side yields,

$$V^{-\rho} = \left(\beta K^{-\rho} + \alpha L^{-\rho}\right). \tag{F.49}$$

Raising both sides by $-\,1/\rho$ yields the full production function,

$$V = \left(\beta K^{-\rho} + \alpha L^{-\rho}\right)^{-1/\rho}. \tag{F.50}$$

The authors have chosen to express this production function in a more symmetrical fashion by

setting $\alpha + \beta = \gamma^{-\rho}$ and $\beta\gamma^{\rho} = \delta$ which yields for $\beta = \dfrac{\delta}{\gamma^{\rho}}$ and for $\alpha = \dfrac{1-\delta\gamma^{\rho}}{\gamma^{\rho}}$ as follows,

$$\alpha + \frac{\delta}{\gamma^{\rho}} = \gamma^{-\rho}. \tag{F.51}$$

Subtracting $\dfrac{\delta}{\gamma^{\rho}}$ from both sides obtaining:

$$\alpha = \gamma^{-\rho} - \frac{\delta}{\gamma^{\rho}}. \tag{F.52}$$

Set the right hand side to have a common fraction,

$$\alpha = \frac{1-\delta}{\gamma^{\rho}}. \tag{F.53}$$

To arrive at the final full production function, substitute the values for α and β into equation (F.50) to obtain,

$$V = \left(\frac{\delta}{\gamma^{\rho}} K^{-\rho} + \frac{1-\delta}{\gamma^{\rho}} L^{-\rho} \right)^{-1/\rho}. \tag{F.54}$$

Raise both sides by $-\rho$ to obtain,

$$V^{-\rho} = \left(\frac{\delta}{\gamma^{\rho}} K^{-\rho} + \frac{1-\delta}{\gamma^{\rho}} L^{-\rho} \right). \tag{F.55}$$

Next factor out the γ^{ρ} term on the right hand side of (F.55) to obtain,

$$V^{-\rho} = \gamma^{-\rho} \left(\delta K^{-\rho} + (1-\delta) L^{-\rho} \right). \tag{F.56}$$

Raising both sides by $-1/\rho$ yields,

$$V = \gamma \left(\delta K^{-\rho} + (1-\delta) L^{-\rho} \right)^{-1/\rho}. \tag{F.57}$$

In most economic literature the ACMS production function utilizes y rather than V as the numéraire expressing equation (F.57) as follows,

$$y = \gamma \left(\delta K^{-\rho} + (1-\delta) L^{-\rho} \right)^{-1/\rho} \tag{F.58}$$

APPENDIX G

CHARACTERISTICS OF PRODUCTION FUNCTIONS

This appendix provides a recapitulation of the production functions noted in Chapter 4 and indicates their salient characteristics.

Table 7. Salient Characteristics of Production Functions

Name	Function	Substitution					RTS	Application	
		Constant/Variable	Perfect σ = ∞	Imperfect 0 < σ < ∞	None σ = 0	Interaction/ Yes or No	Constant/Variable	Macroeconomic	Microeconomic Y= N≥3
Cobb-Douglas	$y = A\prod_{1}^{n} x_i^{\alpha_i}$	C		●		N	C	●	y
Leontief	$y = \min\left(\beta_1 x_1, \beta_2 x_2, \cdots, \beta_n x_n\right)$	C			●	N	C	●	y
Halter, Carter, Hocking Transcendental	$y = A\prod_{1}^{n} x_i^{\alpha_i} e^{\sum_{1}^{n} \gamma_i x_i}$	V		●		N	V	●	y
Heady-Dillon Quadratic	$y = A + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{ij} x_i x_j$	C		●		Y	V	●	y
Heady-Dillon Cubic	$y = a + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \delta_i x_i^2 + \sum_{i=1}^{n} \gamma_i x_i^3$	C		●		Y	V	●	y
Heady-Dillon Square Root	$y = A + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \delta_i x_i^{.5} + .5\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} x_i^{.5} x_j^{.5}$	C		●		Y	V	●	y
Heady-Dillon One and Half Power	$y = A + \sum_{i=1}^{n} \beta_i x_i + \sum_{i=1}^{n} \delta_i x_i^{1.5} + .5\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij} x_i x_j$	C		●		Y	V	●	y
Heady-Dillon Inverse	$y = \frac{1}{A} + \sum_{i=1}^{n} \beta_i^{-1} x_i + .5\sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{ij}^{-1} x_i x_j$	C		●		Y	V	●	y
Newman-Read	$y = A x_1^{\alpha_1} x_2^{\alpha_2} e^{\alpha_3 \ln x_1 \ln x_2}$	C		●		Y	V	●	

Table 7 (cont.)

Name	Function	Substitution					RTS	Application	
		Constant/Variable	Perfect σ = ∞	Imperfect 0 < σ < ∞	None σ = 0	Interaction/ Yes or No	Constant/Variable	Macroeconomic	Microeconomic Y= $N≥3$
Arrow, Chenery, Minhas, Solow (ACMS)	$y = A\left(\beta x_1^{-\alpha} + (1-\beta)x_2^{-\alpha}\right)^{-\frac{1}{\alpha}}$	C		•		N	C	•	
Brown-DeCani	$y = A\left(\beta x_1^{-\alpha} + (1-\beta)x_2^{-\alpha}\right)^{-\frac{v}{\alpha}}$	C		•		N	V	•	
Uzawa-McFadden	$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha}\right)^{-\frac{v}{\alpha}}$	C		•		N	V	•	y
Ferguson-Pfouts (1)	$y = x_1^{\alpha_1} x_2^{1-\alpha_1} e^{\alpha_2 x_1}$	C		•		N	V	•	
Ferguson-Pfouts (2)	$y = x_1^{\alpha_1 - \left(\frac{\alpha_3}{x_2}\right)} x_2^{\alpha_2}$	C		•		N	V	•	
Hildebrand-Liu-Bruno	$y = A\left(\beta_1 x_1^{(1-\alpha)\gamma} x_2^{\alpha\gamma} + (1-\beta)x_2^{\gamma}\right)^{\frac{1}{\gamma}}$	C		•		N	V	•	
Mukerji	$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha_i}\right)^{-\frac{1}{\alpha_0}}$	C		•		N	V	•	y
Ferguson Transcendental	$y = Ax_1^{\alpha_1} x_2^{\alpha_2} e^{\gamma\left(\frac{x_1}{x_2}\right)}$	C		•		N	C	•	
Nerlove-Ringstad (1)	$y = \left(A\prod_{i=1}^{n} x_i^{\alpha_i}\right)^{\frac{1}{\gamma+\beta \ln y}}$	C		•		N	V	•	y
Nerlove-Ringstad (2)	$y = Ae^{\left(-a\prod_{i=1}^{n} x_i^{-\alpha_i}\right)}$	C		•		N	V	•	y
Sato CES (1)	$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho}\right)^{-\frac{1}{\rho}}$ $Z_s = \left(\sum_{i\in s} \gamma_i^s \left(x_i^s\right)^{-\rho_s}\right)^{\frac{-1}{\rho_s}}$	V		•		N	C		y
Sato CES (2)	$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho}\right)^{-\frac{v}{\rho}}$	V		•		N	V	•	y
Sato CES (3)	$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho_s'}\right)^{-\frac{v}{\rho}}$	V		•		N	V	•	y

Table 7 (cont.)

Name	Function	Substitution					RTS	Application	
		Constant/Variable	Perfect σ = ∞	Imperfect 0 < σ < ∞	None σ = 0	Interaction/ Yes or No	Constant/Variable	Macroeconomic	Microeconomic Y= N≥3
Revankar Generalized (1)	$y = Ax_1^{(1-\alpha\gamma)}\left(x_2+(\gamma-1)x_1\right)^{\alpha\gamma}$	V		•		N	C	•	
Revankar Generalized (2)	$y = Ax_1^{v(1-\alpha\gamma)}\left(x_2+(\gamma-1)x_1\right)^{v\alpha\gamma}$	V		•		N	V	•	
Kmenta	$y = ae^{-.5(1-\alpha)\alpha\gamma v(\ln x_1-\ln x_2)^2+\alpha v\ln x_1+(1-\alpha)v\ln x_2}$	C		•		N	V	•	
Bruno	$y = Ax_1^{\alpha}x_2^{1-\alpha}-mx_1$	C		•		N	C	•	
Lu-Fletcher	$y = A\left(\beta x_1^{-\rho}+(1-\beta)\gamma\left(\frac{x_1}{x_2}\right)^{\delta(1-\rho)}x_2^{-\rho}\right)^{-\frac{1}{\rho}}$	V		•		N	C	•	
Sato-Hoffman VES1	$y = A\left(\frac{x_2}{x_1}\right)^{\beta}e^{\alpha\left(x_2/x_1\right)}$	V		•		N	C	•	
Sato-Hoffman VES2	$y = Ax_2^{\alpha\left(1/1+\gamma\right)}\left(x_1+\left(\beta x_2/1+\gamma\right)x_1\right)^{\alpha\left(\gamma/1+\gamma\right)}$	V		•		N	V	•	
Sato-Hoffman VES3	$y = A\left(\delta x_1^{\sigma-\left(1/\sigma\right)}+(1-\delta)x_2^{\sigma-\left(1/\sigma\right)}\right)^{\sigma/(\sigma-1)}$	V		•		N	C	•	
Sato CEDD (1)	$y = \left(\alpha_1/1-\delta_1\right)x_1^{\delta_1}x_2^{(1-\delta_1)}+\beta_1x_1$	V		•		N	C	•	
Sato CEDD (2)	$y = \left(\alpha_2/1-\delta_2\right)x_2^{\delta_2}x_1^{(1-\delta_2)}+\beta_2x_2$	V		•		N	C	•	
Sato CEDD (3)	$y = \alpha_1x_1\ln\left(\frac{x_2}{x_1}\right)+\beta_1x_1$	V		•		N	V	•	
Sato CEDD (4)	$y = \alpha_2x_2\ln\left(\frac{x_1}{x_2}\right)+\beta_2x_2$	V		•		N	V	•	
Chu, Aigner, and Frankel Log Quadratic	$y = A\left(x_1/\overline{x}_1\right)^{\alpha_1\left((1-\ln x_1)/\ln\overline{x}_1\right)}\left(x_2/\overline{x}_2\right)^{\alpha_2\left((1-\ln x_2)/\ln\overline{x}_2\right)}$	V		•		N	V	•	
Vazquez (1)	$y = X_2^{\left(1-1/\beta\right)}X_1^{\left(1/\beta\right)}+\left(\alpha/1-\beta\right)X_1+\gamma X_2$	V		•		N	C	•	
Vazquez (2)	$y = X_2^{\left(1-1/\beta\right)}X_1^{\left(1/\beta\right)}+\gamma X_2$	V		•		N	C	•	

Table 7 (cont.)

Name	Function	Substitution					RTS	Application	
		Constant/Variable	Perfect σ = ∞	Imperfect 0 < σ < ∞	None σ = 0	Interaction/ Yes or No	Constant/Variable	Macroeconomic	Microeconomic Y= N≥3
Vazquez (3)	$y = A\left(\beta x_1^{-\gamma/(1+\alpha)} + (1+\alpha)\left(\frac{x_2}{x_1}\right)^{-\alpha\gamma} x_2^{-\gamma/(1+\alpha)}\right)^{-\lambda/\gamma(1+\alpha)}$	V		•		N	V	•	
Sargan/CJL	$\ln y = \ln A + \sum_{i=1}^{n} \alpha_i \ln x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j$	V		•		N	V	•	y
Diewert Generalized Leontief	$y = \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij} x_i^{1/2} x_j^{1/2}$	V		•		N	C	•	y
Kadiyala (1)	$y = A\left(\alpha_{11} x_1^{2\beta_0} + 2\alpha_{12} x_1^{\beta_1} x_2^{\beta_2} + \alpha_{22} x_2^{2\beta_0}\right)$	V		•		Y	C	•	
Kadiyala (2)	$y = A\left(\sum_{i=1}^{n} \alpha_{ii} x_i^{2\beta} + 2\sum_{i<j}^{n} \alpha_{ij} x_i^{\beta} x_j^{\beta}\right)$	V		•		Y	C	•	y
Vinod (1)	$y = e^{\alpha_0} x_1^{\alpha_1 + \alpha_3 \ln x_2} x_2^{\alpha_2}$	V		•		Y	V	•	
Vinod (2)	$\ln y = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + \sum_{i<j}^{n} \alpha_{ij} \ln x_i \ln x_j$	V		•		Y	V	•	y
Lovell (1)	$y = Ae^{\lambda t}\left(\delta x_1^{-\rho} + (1-\delta) x_2^{-\rho}\right)^{-1/\rho}$	C		•		N	C	•	
Lovell (2)	$y = Ae^{\lambda t}\left((1+\beta) x_1 x_2^{\beta} + \alpha x_2^{1+\beta}\right)^{1/1+\beta}$	V		•		Y	C	•	
Lovell (3)	$y = Ae^{\lambda t} x_1^{\alpha} x_2^{1-\alpha} e^{\beta k}$	V		•		N	C	•	
Diewert Generalized Cobb-Douglas	$y = A\prod_{i=1}^{n}\prod_{j=1}^{n}\left(\frac{1}{2}x_i + \frac{1}{2}x_j\right)^{\alpha_{ij}}$	V		•		N	C	•	Y
Denny Generalized Quadratic	$y = \left(\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_{ij} x_i^{\beta\gamma} x_j^{\beta(1-\gamma)}\right)^{\mu/\beta}$	C / V		•		Y	C/V	•	y
Helmy GTPF (1)	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} x_i x_j}$	V		•		Y	V	•	y
Helmy GTPF(2)	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} \ln x_i \ln x_j}$	V		•		Y	V	•	Y

Table 7 (cont.)

Name	Function	Substitution					RTS	Application	
		Constant/Variable	Perfect σ = ∞	Imperfect 0 < σ < ∞	None σ = 0	Interaction/ Yes or No	Constant/Variable	Macroeconomic	Microeconomic Y= N≥3
Helmy GTPF(3)	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i x_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} x_i^{1/2} x_j^{1/2}}$	V		•		Y	V	•	y
Helmy Quadratic Logarithmic	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n}\beta_i (\ln x_i)^2}$	V		•		Y	V	•	y

APPENDIX H

CHOICE OF FUNCTIONAL FORM

Table 8. Evaluation of Production Functions

Name	Function	Selection Criteria: Application & Objective	Maintained Hypotheses	Parsimony in Parameters	Ease of Interpretation	Computational Ease
Cobb-Douglas	$y = A\prod_{1}^{n} x_i^{\alpha_i}$	2	1	5	5	5
Leontief	$y = \min\left(\beta_1 x_1, \beta_2 x_2, \cdots, \beta_n x_n\right)$	1	1	5	5	5
Halter, Carter, Hocking Transcendental	$y = A\prod_{1}^{n} x_i^{\alpha_i} e^{\sum_{1}^{n}\gamma_i x_i}$			E		
Heady-Dillon Quadratic	$y = A + \sum_{i=1}^{n}\beta_i x_i + \sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij} x_i x_j$			E		
Heady-Dillon Cubic	$y = a + \sum_{i=1}^{n}\beta_i x_i + \sum_{i=1}^{n}\delta_i x_i^2 + \sum_{i=1}^{n}\gamma_i x_i^3$			E		
Heady-Dillon Square Root	$y = A + \sum_{i=1}^{n}\beta_i x_i + \sum_{i=1}^{n}\delta_i x_i^{.5} + .5\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij} x_i^{.5} x_j^{.5}$			E		
Heady-Dillon One and Half Power	$y = A + \sum_{i=1}^{n}\beta_i x_i + \sum_{i=1}^{n}\delta_i x_i^{1.5} + .5\sum_{i=1}^{n}\sum_{j=1}^{n}\gamma_{ij} x_i x_j$			E		
Heady-Dillon Inverse	$y = \frac{1}{A} + \sum_{i=1}^{n}\beta_i^{-1} x_i + .5\sum_{i=1}^{n}\sum_{j=1}^{n}\beta_{ij}^{-1} x_i x_j$			E		
Newman-Read	$y = Ax_1^{\alpha_1} x_2^{\alpha_2} e^{\alpha_3 \ln x_1 \ln x_2}$	E				
Arrow, Chenery, Minhas, Solow (ACMS)	$y = A\left(\beta x_1^{-\alpha} + \left(1-\beta\right) x_2^{-\alpha}\right)^{-\frac{1}{\alpha}}$	E				
Brown-DeCani	$y = A\left(\beta x_1^{-\alpha} + (1-\beta) x_2^{-\alpha}\right)^{-\frac{\nu}{\alpha}}$	E				

Table 8 (cont.)

Name	Function	Selection Criteria				
		Application & Objective	Maintained Hypotheses	Parsimony in Parameters	Ease of Interpretation	Computational Ease
Uzawa-McFadden	$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha}\right)^{-\frac{v}{\alpha}}$	3	3	4	5	1
Ferguson-Pfouts (1)	$y = x_1^{\alpha_1} x_2^{1-\alpha_1} e^{\alpha_2 x_1}$	**E**				
Ferguson-Pfouts (2)	$y = x_1^{\alpha_1 - \left(\frac{\alpha_3}{x_2}\right)} x_2^{\alpha_2}$	**E**				
Hildebrand-Liu-Bruno	$y = A\left(\beta_1 x_1^{(1-\alpha)\gamma} x_2^{\alpha\gamma} + (1-\beta) x_2^{\gamma}\right)^{\frac{1}{\gamma}}$	**E**				
Mukerji	$y = A\left(\sum_{i=1}^{n} \beta_i x_i^{-\alpha_i}\right)^{-\frac{1}{\alpha_0}}$	3	4	3	4	1
Ferguson Transcendental	$y = Ax_1^{\alpha_1} x_2^{\alpha_2} e^{\gamma\left(\frac{x_1}{x_2}\right)}$	**E**				
Nerlove-Ringstad (1)	$y = \left(A\prod_{i=1}^{n} x_i^{\alpha_i}\right)^{\frac{1}{\gamma+\beta\ln y}}$	3	3	4	1	1
Nerlove-Ringstad (2)	$y = Ae^{\left(-a\prod_{i=1}^{n} x_i^{\alpha_i}\right)}$			E		
Sato CES (1)	$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho}\right)^{-\frac{1}{\rho}}$ $Z_s = \left(\sum_{i \in s} \gamma_i^s \left(x_i^s\right)^{-\rho_s}\right)^{\frac{-1}{\rho_s}}$			**E**		
Sato CES (2)	$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho}\right)^{-\frac{v}{\rho}}$			**E**		
Sato CES (3)	$y = \left(\sum_{s=1}^{s} \beta_s Z_s^{-\rho_s'}\right)^{-\frac{v}{\rho}}$			**E**		
Revankar Generalized (1)	$y = Ax_1^{(1-\alpha\gamma)}\left(x_2 + (\gamma-1)x_1\right)^{\alpha\gamma}$	**E**				
Revankar Generalized (2)	$y = Ax_1^{v(1-\alpha\gamma)}\left(x_2 + (\gamma-1)x_1\right)^{v\alpha\gamma}$	**E**				
Kmenta	$y = ae^{-.5(1-\alpha)\alpha\gamma v(\ln x_1 - \ln x_2)^2 + \alpha v \ln x_1 + (1-\alpha) v \ln x_2}$	**E**				

Table 8 (cont.)

Name	Function	Selection Criteria: Application & Objective	Maintained Hypotheses	Parsimony in Parameters	Ease of Interpretation	Computational Ease
Bruno	$y = Ax_1^{\alpha} x_2^{1-\alpha} - mx_1$	**E**				
Lu-Fletcher	$y = A\left(\beta x_1^{-\rho} + (1-\beta)\gamma \left(\frac{x_1}{x_2} \right)^{\delta(1-\rho)} x_2^{-\rho} \right)^{-\frac{1}{\rho}}$	**E**				
Sato-Hoffman VES1	$y = A\left(\frac{x_2}{x_1} \right)^{\beta} e^{\alpha\left(x_2 / x_1 \right)}$	**E**				
Sato-Hoffman VES2	$y = Ax_2^{\alpha\left(1/(1+\gamma) \right)} \left(x_1 + \left(\beta x_2 / (1+\gamma) \right) x_1 \right)^{\alpha\left(\gamma/(1+\gamma) \right)}$	**E**				
Sato-Hoffman VES3	$y = A\left(\delta x_1^{\sigma-\left(1/\sigma \right)} + (1-\delta) x_2^{\sigma-\left(1/\sigma \right)} \right)^{\sigma/(\sigma-1)}$	**E**				
Sato CEDD (1)	$y = \left(\alpha_1 / (1-\delta_1) \right) x_1^{\delta_1} x_2^{(1-\delta_1)} + \beta_1 x_1$	**E**				
Sato CEDD (2)	$y = \left(\alpha_2 / (1-\delta_2) \right) x_2^{\delta_2} x_1^{(1-\delta_2)} + \beta_2 x_2$	**E**				
Sato CEDD (3)	$y = \alpha_1 x_1 \ln\left(\frac{x_2}{x_1} \right) + \beta_1 x_1$	**E**				
Sato CEDD (4)	$y = \alpha_2 x_2 \ln\left(\frac{x_1}{x_2} \right) + \beta_2 x_2$	**E**				
Chu, Aigner, and Frankel Log Quadratic	$y = A\left(x_1 / \overline{x}_1 \right)^{\alpha_1\left((1-\ln x_1) / \ln \overline{x}_1 \right)} \left(x_2 / \overline{x}_2 \right)^{\alpha_2\left((1-\ln x_2) / \ln \overline{x}_2 \right)}$	**E**				
Vazquez (1)	$y = X_2^{\left(1-1/\beta \right)} X_1^{\left(1/\beta \right)} + \left(\alpha / (1-\beta) \right) X_1 + \gamma X_2$	**E**				
Vazquez (2)	$y = X_2^{\left(1-1/\beta \right)} X_1^{\left(1/\beta \right)} + \gamma X_2$	**E**				
Vazquez (3)	$y = A\left(\beta x_1^{-\gamma/(1+\alpha)} + (1+\alpha)\left(\frac{x_2}{x_1} \right)^{-\alpha\gamma} x_2^{-\gamma/(1+\alpha)} \right)^{-\lambda/\gamma(1+\alpha)}$	**E**				

Table 8 (cont.)

Name	Function	Selection Criteria				
		Application & Objective	Maintained Hypotheses	Parsimony in Parameters	Ease of Interpretation	Computational Ease
Sargan/CJL	$\ln y = \ln A + \sum_{i=1}^{n} \alpha_i \ln x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j$			**E**		
Diewert Generalized Leontief	$y = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} x_i^{1/2} x_j^{1/2}$			**E**		
Kadiyala (1)	$y = A\left(\alpha_{11} x_1^{2\beta_0} + 2\alpha_{12} x_1^{\beta_1} x_2^{\beta_2} + \alpha_{22} x_2^{2\beta_0}\right)$	**E**				
Kadiyala (2)	$y = A\left(\sum_{i=1}^{n} \alpha_{ii} x_i^{2\beta} + 2\sum_{i<j}^{n} \alpha_{ij} x_i^{\beta} x_j^{\beta}\right)$			**E**		
Vinod (1)	$y = e^{\alpha_0} x_1^{\alpha_1 + \alpha_3 \ln x_2} x_2^{\alpha_2}$			**E**		
Vinod (2)	$\ln y = \alpha_0 + \sum_{i=1}^{n} \alpha_i \ln x_i + \sum_{i<j}^{n} \alpha_{ij} \ln x_i \ln x_j$			**E**		
Lovell (1)	$y = Ae^{\lambda t}\left(\delta x_1^{-\rho} + (1-\delta) x_2^{-\rho}\right)^{-1/\rho}$	**E**				
Lovell (2)	$y = Ae^{\lambda t}\left((1+\beta) x_1 x_2^{\beta} + \alpha x_2^{1+\beta}\right)^{1/1+\beta}$	**E**				
Lovell (3)	$y = Ae^{\lambda t} x_1^{\alpha} x_2^{1-\alpha} e^{\beta k}$	**E**				
Diewert Generalized Cobb-Douglas	$y = A\prod_{i=1}^{n} \prod_{j=1}^{n} \left(\frac{1}{2} x_i + \frac{1}{2} x_j\right)^{\alpha_{ij}}$			**E**		
Denny Generalized Quadratic	$y = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_{ij} x_i^{\beta\gamma} x_j^{\beta(1-\gamma)}\right)^{\mu/\beta}$			**E**		
Helmy GTPF (1)	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n} \beta_i x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} x_i x_j}$			**E**		
Helmy GTPF(2)	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n} \beta_i x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} \ln x_i \ln x_j}$			**E**		
Helmy GTPF(3)	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n} \beta_i x_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \beta_{ij} x_i^{1/2} x_j^{1/2}}$			**E**		
Helmy Quadratic Logarithmic	$y = A\prod_{i=1}^{n} x_i^{\alpha_i} e^{\sum_{i=1}^{n} \beta_i (\ln x_i)^2}$			**E**		

BIBLIOGRAPHY

Aigner, D. J., C. A. K. Lovell, and P Schmidt. "Formulation and Estimation of Stochastic Frontier Production Functions." *Journal of Econometrics* 6 (1977): 21-37.

American Institute of Architects. *AIA Firm Survey 2000-2002*. Washington, D.C.: American Institute of Architects, 2003.

———. *The Architect's Handbook of Professional Practice*. Fourteenth ed. Washington, D.C.: John Wiley & Sons, Inc., 2008.

———. *Architecture Factbook*. Washington, D.C.: American Institute of Architects, 1994.

———. *The Business of Architecture*. Washington, D.C.: American Institute of Architects, 2006.

———. *The Business of Architecture*. Washington, D.C.: American Institute of Architects, 2009.

———. *Compsensation Guidelines for Architectural and Engineering Services: A Management Guide to Cost-Based Compensation*. Second ed. Washington, D.C.: American Institute of Architects, 1978.

———. *Financial Management for Achitectural Firms*. Washington, D.C.: American Institute of Architects, 1970.

———. *Instructions: Standard Accounting for Architects*. Washington, D.C.: American Institute of Architects, 1949.

———. *Manual of Accounting for Architects*. Los Angeles, California: Parker, Stone & Baird Co., 1935.

———. *Standard Accounting for Architects: A Management Guide to Accounting and Financial Management Procedures*. Washington, D.C.: American Institute of Architects, 1978.

Arrow, Kenneth J., Hollis B. Chenery, Bagicha S. Minhas, and Robert M. Solow. "Capital-Labor Substitution and Economic Efficiency." *The Review of Economics and Statistics* 43, no. 3 (1961): 225-50.

Aspromourgos, Tony. "On the Origins of the Term 'Neoclassial'." *Cambridge Journal of Economics* 10, no. 3 (1986): 265-70.

Baker, Kermit, and Pradeep Dalal. *1997 AIA Firm Survey Report*. Washington, D.C.: American Institute of Architects, 1997.

Baker, Paul R. *Richard Morris Hunt*. Cambridge, Massachusetts: The MIT Press, 1980.

Bartlett, John. *Familiar Quotations: A Collection of Passages, Phrases and Proverbs Traced to Their Sources in Ancient and Modern Literature*. Boston, Massachusetts: Little, Brown and Company, 1980.

Baumol, William J., and Stephen M. Goldfeld. "Precursors in Mathematical Economics: An Anthology." London, England: The London School of Economics and Political Science (University of London), 1968.

Beattie, Bruce R., and C. Robert Taylor. *The Economics of Production*. New York, New York: John Wiley & Sons, 1985.

Beightler, Charles S., Don T. Phillips, and Douglass J. Wilde. *Foundation of Optimization*. Second ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1979.

Berndt, Ernst R., and Laurits R. Christensen. "The Translog Function and the Substitution of Equipment, Structures, and Labor in U.S. Manufacturing 1929-68." *Journal of Econometrics* 1, no. 1 (1973): 81-113.

Blackorby, Charles, Daniel Primont, and R. Robert Russell. *Duality, Separability, and Functional Structure: Theory and Economic Applications*. New York, New York: Elsevier North-Holland, 1978.

Boggs, Paul T., and Jon W. Tolle. "Sequential Quadratic Programming." *Acta Numerica* 4 (Jan 1995): 1-51.

Bosworth, Derek L. *Production Functions: A Theoretical and Empirical Study*. Lexington, Massachusetts: Saxon House, D.C. Heath Ltd. and Lexington books, D.C. Heath & Co., 1976.

Brown, Murray, ed. *The Theory and Empirical Analysis of Production*. Vol. 31, Studies in Income and Wealth. New York, New York: National Bureau of Economic Research and Columbia University Press, 1967.

Brown, Murray, and John S. De Cani. "Technological Change and the Distribution of Income." *International Economic Review* 4, no. 3 (1963): 289-309.

Bruno, Michael. "Estimation of Factor Contribution to Growth under Structural Disequilibrium." *International Economic Review* 9, no. 1 (1968): 49-62.

———. *A Note of the Implications of an Empirical Relationship between Output Per Unit of Labor, the Wage Rate, and the Capital Labor Ratio*. Stanford, California: Stanford University, 1962.

Carlson, Sune. *A Study on the Pure Theory of Production*. New York, New York: Augustus M. Kelly, 1965.

Case and Company Inc. *The Economics of Architectural Practice*. Washington, D.C.: American Institute of Architects, 1968.

Case and Company Inc. *Profit Planning in Architectural Practice*. Washington, D.C.: American Institute of Architects, 1968.

Chambers, Robert G. *Applied Production Analysis*. Cambridge, England: Cambridge University Press, 1988.

Ching, Chauncey T. K., and John F. Yanagida. *Production Economics: Mathematical Development and Applications*. New Brunswick, New Jersey: Transaction Books, 1985.

Christensen, Laurits R., Dale W. Jorgenson, and Lawrence J. Lau. "Conjugate Dualtiy and the Transcendental Logarithmic Production Function." *Econometrica* 39, no. 4 (1971): 255-56.

Chu, S. F., D. J. Aigner, and M. Frankel. "On the Log-Quadratic Law of Production." *Southern Economic Journal* 37, no. 1 (1970): 32-39.

Clark, John Bates. *The Distribution of Wealth: A Theory of Wages, Interest and Profits*. New York: Macmillan, 1899.

Coase, Ronald H. "The Nature of the Firm." *Economica* 4, no. 16 (1937): 386-405.

Cobb, Charles W., and Paul H. Douglas. "A Theory of Production." *The American Economic Review* 18, no. 1 (1928): 139-65.

Corbo, Vittorio, and Patricio Meller. "The Translog Production Function." *Journal of Econometrics* 10, no. 2 (1979): 193-99.

Dempsey, Bernard W. *The Frontier Wage: The Economic Organization of Free Agents*. Chicago, Illinois: Loyola University Press, 1960.

Denny, Michael G. S. "The Relationship between Functional Forms for the Production System." *The Canadian Journal of Economics* 7, no. 1 (Feb 1974): 21-31.

Dharmapala, Dhammika, and Michael McAleer. "Econometric Methodology and the Philosophy of Science." *Journal of Statistical Planning and Inference* 49, no. 1 (1996): 9-37.

Diewert, W. Erwin. "An Application of the Shepard Duality Theorem: A Generalized Leontief Production Function." *The Journal of Political Economy* 79, no. 3 (May-Jun 1971): 481-507.

———. "Separability and a Generalization of the Cobb-Douglas Cost, Production and Indirect Utility Functions." 1-45. Vancouver, British Columbia, Canada: Stanford University and University of British Columbia, 1973.

Doll, John P., and Frank Orazem. *Production Economics: Theory with Applications*. New York, New York: Joh Wiley & Sons, 1984.

Douglas, Paul H. "Are There Laws of Production?". *The American Economic Review* 38, no. 1 (Mar 1948): 1-41.

———. *The Theory of Wages*. New York, New York: Kelly & Millman, Inc., 1957.

Durand, David. "Some Thoughts on Marginal Productivity, with Special Reference to Professor Douglas' Analysis." *Journal of Political Economy* 45, no. 6 (1937): 740-58.

Erdos, Paul, and Underwood Dudley. "Some Remarks and Problems in Number Theory Related to the Work of Euler." *Mathematics Magazine* 56, no. 5 (1983): 292-98.

Farrell, M. J. "The Measurement of Productive Efficiency." *Journal of the Royal Statistical Society* 120 (1957): 253-81.

Ferguson, C. E. "Capital-Labor Substitution and Technological Progress in the U.S.: Statistical Evidence from a Transcendental Production Function." Memo, 1965.

———. *The Neoclassical Theory of Production and Distribution*. London, England: Cambridge University Press, 1969.

Ferguson, C. E., and Ralph W. Pfouts. "Aggregate Production Functions and Relative Factor Shares." *International Economic Review* 3, no. 3 (1962): 328-37.

Fisher, Thomas. "Who Makes What and How We Might All Make More." *Progressive Architecture*, no. Dec (1995): 49-53,96.

Fuss, Melvyn, and Daniel McFadden. *Production Economics: A Dual Approach to Theory and Applications, Vol I and Ii*. New York, New York: North-Holland Publishing Company, 1978.

Getz, Lowell, and Frank Stasiowski. *Financial Management for the Design Professional a Handbook for Architects, Engineers, and Interior Designers*. New York, New York: Whitney Library of Design, 1984.

Gherity, James A., ed. *Econoic Thought: A Historical Anthology*. New York, New York: Random House, 1965.

Gossen, Hermann Heinrich. *The Laws of Human Relations and the Rules of Human Action Derived Therefrom*. Translated by Rudopph C. Blitz. Cambridge, Massachusetts: The

MIT Press, 1983. Entwickelung der Gesetze des menschlichen Verkehrs, und der daraus fliessenden Regeln fur menschliches Handeln.

Griliches, Zvi, and Vidar Ringstad. *Economies of Scale and the Form of the Production Function: An Econometric Study of Norwegian Manufacturing Establishment Data*. Contributions to Economic Analysis. edited by J. Johnston, D. W. Jorgenson and J. Waelbroeck. Vol. 72, Amsterdam, The Netherlands: North-Holland Publishing Company, 1971.

Halter, A. N., H. O. Carter, and J. G. Hocking. "A Note on the Transcendental Production Function." *Journal of Farm Economics* 39, no. 4 (1957): 966-74.

Heady, Earl O., and John L. Dillon. *Agricultural Production Functions*. Ames, Iowa: Iowa State University Press, 1961.

Helmy, Aly A. "A Family of Generalized Transcendental Production Functions." Dissertation, University of Nore Dame, 1981.

Henderson, James M., and Richard E. Quandt. *Microeconomic Theory: A Mathematical Approach*. Third ed. New York, New York: McGraw-Hill Book Company, 1980.

Hicks, Douglas T. *Activity-Based Costing: Making It Work for Small and Mid-Sized Companies*. Second ed. New York, New York: John Wiley & Sons, Inc., 1999.

Hildebrand, George H., and Dazhong Liu. *Manufacturing Production Functions in the United States*. Cornell Studies in Industrial and Labor Relations. Vol. 15, Ithaca, New York: New York State School of Industrial and Labor Relations, Cornell University, 1965.

Humphrey, Thomas M. "Algebraic Production Functions and Their Uses before Cobb-Douglas." *Economic Quarterly* 83, no. 1 (Winter 1997): 51-83.

Hyman, David N. *Modern Microeconomics: Analysis and Applications*. St. Louis, Missouri: Times Mirror/Mosby College Publishing, 1986.

Jevons, William Stanley. *The Theory of Political Economy*. Harmondsworth, Middlesex, England: Penguin Books Ltd., 1871.

Kadiyala, Koteswara Rao. "Production Functions and Elasticity of Substitution." *Southern Economic Journal* 38, no. 3 (Jan 1972): 281-84.

Kamerschen, David R., and Lloyd M. Valentine. *Intermediate Microeconomic Theory*. Cincinnati, Ohio: South-Western Publishing, 1977.

Kaplan, Robert S., and Robin Cooper. *Cost & Effect: Using Integrated Cost Systems to Drive Profitability and Performance*. Boston, Massachusetts: Harvard Business School Press, 1998.

Kaplan, Robert S., and H. Thomas Johnson. *Relevance Lost: The Rise and Fall of Management Accounting*. Boston, Massachusetts: Harvard Business School Press, 1987.

Kmenta, Jan. "On Estimation of the CES Production Function." *International Economic Review* 8, no. 2 (1967): 180-89.

Kubany, Elizabeth Harrison, and Charles D. Linn. "The Fee Dilemma Part I: Why Architects Don't Charge Enough." *Architectural Record* 187, no. 10 (Oct 1999): 110-17.

———. "The Fee Dilemma Part II: How to Increase Your Fees in a Tough Market." *Architectural Record* 187, no. 11 (Nov 1999): 60-66.

Layard, Richard, J. Denis Sargan, Margaret E. Ager, and Deborah J. Jones. *Qualified Manpower and Economic Performance: An Inter-Plant Study in the Electrical Engineering Industry*. London, England: Allen Lane The Penguin Press, 1971.

Leibenstein, Harvey. "Allocative Efficiency Vs. 'X-Efficiency." *The American Economic Review* 56, no. 3 (1966): 392-415.

Leontief, Wassily W. *Essays in Economics: Theories and Theorizing*. London, England: Oxford University Press, 1966.

———. *Input-Output Economics*. Oxford, England: Oxford University Press, 1986.

———. *The Structure of American Economy, 1919-1929: An Empirical Application of Equilibrium*. Cambridge, Massachusetts: Harvard University Press, 1941.

Lewis, Roger K. *Architect?: A Candid Guide to the Profession*. Cambridge, Massachusetts: The MIT Press, 1998.

Loftsguard, Laurel D., and Earl O. Heady. "Application of Dynamic Programming Models for Optimum Farm and Home Plans." *Journal of Farm Economics* 41 (1959): 51-62.

Lovell, C. A. Knox. "Estimation and Prediction with CES and VES Production Functions." *International Economic Review* 14, no. 3 (Oct 1973): 676-92.

Lu, Yao-chi, and Lehman B. Fletcher. "A Generalization of the CES Production Function." *The Review of Economics and Statistics* 50, no. 4 (1968): 449-52.

Lusht, Kenneth M. *Real Estate Valuation: Principles and Applications*. State College, PA: KML Publishing, 2001.

Marshall, Alfred. *Principles of Economics*. Reprint of the Eighth ed. Amherst, New York: Prometheus Books, 1997. Principles of Economics.

———. *Principles of Economics*. Vol. 1, Cambridge, England: Cambridge University Press, 1890.

Mattox, Robert F. *Standardized Accounting for for Architects: A Guide to Understanding, Developing and Implementing Accounting Procedures*. Third ed. Washington, D.C.: American Institue of Architects, 1982.

McFadden, Daniel. "Constant Elasticity of Substitution Production Functions." *The Review of Economic Studies* 30, no. 2 (1963): 73-83.

Moncur, Michael. "The Quotation Page." www.quotationspage.com/quotes/George_Bernard_Shaw.

Morgan, Jim. *Management for the Small Design Firm: Handling Your Practice, Personnel, Finances, and Projects*. New York, New York: Whitney Library of Design, 1998.

Morrison-Paul, Catherine J. *Cost Structure and the Measurement of Economic Performance*. Boston, Massachusetts: Kluwer Academic Publishers, 1999.

Mukerji, V. "A Generalized S.M.A.C. Function with Constant Ratios of Elasticity of Substitution." *The Review of Economic Studies* 30, no. 3 (Oct 1963): 233-36.

Negishi, Takashi. *History of Economic Theory*. Amsterdam, The Netherlands: Elsevier Science Publishers B.V., 1989.

Nerlove, Marc. "Returns to Scale in Electricity Supply." In *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld*, edited by Carl Christ. Standford, California: Stanford University Press, 1963.

Newman, P. K., and R. C. Read. "Production Functions with Restricted Input Shares." *International Economic Review* 2, no. 1 (1961): 127-33.

Niehans, Jurg. *A History of Economic Theory: Classic Contributions, 1720-1980*. Baltimore, Maryland: The Johns Hopkins University Press, 1990.

Nobel Media AB. http://nobelprize.org/nobel_prizes/medicine/laureates/.

Oser, Jacob, and Stanley L. Brue. *The Evolution of Economic Thought*. Fourth ed. SanDiego, California: Harcourt Brace Jovanovich, Inc, 1988.

Pappas, James L., and Eugene F. Brigham. *Managerial Economics*. Third ed. Hinsdale, Illinois: The Dryden Press, 1979.

Petersen, H. Craig, and W. Cris Lewis. *Managerial Economics*. Third ed. New York, New York: Macmillan Publishing Company, 1994.

Revankar, Nagesh S. "Production Functions with Variable Elasticity of Substitution and Variable Returns to Scale." Dissertation, University of Wisconsin, 1967.

Ringstad, Vidar. "Econometric Analyses Based on a Production Function with Neutrally Variable Scale-Elasticity." *Swedish Journal of Economics* 69, no. 2 (June 1967): 115-33.

Samuelson, Paul A. *Foundations of Economic Analysis*. New York, New York: Atheneum, 1947.

Sato, K. "A Two-Level Constant-Elasticity-of-Substitution Production Function." *The Review of Economic Studies* 34, no. 2 (1967): 201-18.

Sato, Ryuzo. "The Estimation of Biased Technical Progress and the Production Function." *International Economic Review* 11, no. 2 (1970): 179-208.

Sato, Ryuzo, and Ronald F. Hoffman. "Production Functons with Variable Elasticity of Factor Substitution: Some Analysis and Testing." *The Review of Economics and Statistics* 50, no. 4 (1968): 453-60.

Schumpeter, Joseph A. *History of Economic Analysis*. New York, New York: Oxford University Press, 1954.

Shephard, Ronald W. *Cost and Production Functions*. Princeton, New Jersey: Princetion University Press, 1953.

Smith, Adam. *An Inquiry into the Nature and Causes of the Wealth of Nations*. London, England: Strahan and Cadell, 1776.

Stefanou, Sprio E., and Kristiaan Kerstens. "Applied Production Analysis Unveiled in Open Peer Review: Introductory Remarks." *Journal of Productivity Analysis* 30, no. 1 (2008): 1-6.

Stigler, George J. *Production and Distribution Theories: The Formative Period*. New York, New York: Agathon Press, Inc., 1968.

Sydsaeter, Knut, and Peter J. Hammond. *Mathematics for Economic Analysis*. Englewood Cliffs, New Jersey: Prentice Hall, 1995.

Thinkexist.com. "Thinkexist.Com." www.thinkexist.com/quotes/john_kenneth_galbraith.

Thompson, Arthur A. *Economics of the Firm: Theory and Practice*. Third ed. Englewood Cliffs, New Jersey: Prentice-Hall, 1981.

Thompson, Gary D. "Choice of Flexible Functional Forms: Review and Appraisal." *Western Journal of Agricultural Economics* 13, no. 2 (1988): 169-83.

Tullock, Gordon. "A Plea to Economists Who Favor Liberty." *Eastern Economic Journal* 27, no. 2 (Spring 2001): 203-07.

Uzawa, Hirofumi. "Production Functions with Constant Elasticities of Substitution." *The Review of Economic Studies* 29, no. 4 (1962): 291-99.

Varian, Hal R. *Intermediated Microeconomics: A Modern Approach*. Seventh ed. New York, New York: W. W. Norton & Company, 2006.

Vazquez, Andres. "Homogeneous Production Functions with Constant or Variable Elasticity of Substitution." *Journal of Institutional and Theoretical Economics* 127, no. 1 (1971): 7-26.

Veblen, Thorstein. "The Preconceptions of Economic Science." *The Quarterly Journal of Economics* 14, no. 2 (Feb 1900): 240-69.

Vinod, Hrishikesh D. "Nonhomogeneous Production Functions and Applications to Telecommunications." *The Bell Journal of Economics and Management Science* 3, no. 2 (Autumn 1972): 531-43.

Von Thunen, Johann Heinrich. *Von Thunen's Isolated State*. Translated by Carla M. Wartenberg. Oxford, England: Pergamon Press, 1966. Der Isolierte Staat.

Walras, Leon. *Elements of Pure Economics*. Translated by Willian Jaffe. The American Economic Association and The Royal Economic Society by Richard D. Irwin, Inc., 1926.

Walters, Alan Arthur. "Production and Cost Functions: An Econometric Survey." *Econometrica* 31, no. 1/2 (Jan-Apr 1963): 1-66.

Warren, Carl S., Philip E. Fess, and James M. Reeve. *Accounting*. 18th ed. Cincinnati, Ohio: South-Western College Publishing, 1996.

Wicksteed, Philip H. *An Essay on the Co-Ordination of the Laws of Distribution*. London, England: Macmillan & Co., 1894.

Wintner, Steve L., and Michael Tardif. *Financial Management for Design Professionals: The Path to Profitability*. Chicago, Illinois: Kaplan AEC Education, 2007.

Zellner, A., and N. S. Ravenkar. "Generalized Production Functions." *Review of Economic Studies* 36, no. 106 (1969): 241-50.