CHAPTER 1:

INTRODUCTION

This book is about the career processes and outcomes of women in science. The brevity of this statement belies the complexity of the topic and the scope of the study that we present in the pages that follow. The underrepresentation of women is a hallmark characteristic of science and has persisted in spite of the rapid improvement of women's social and economic status over the past few decades. During this period the gender gap in overall educational attainment in the U.S. has closed (Bae et al. 2000), and women have made significant inroads into many other high status professions, such as medicine, the law, and the arts (Reskin and Roos 1990). In trend analyses of occupational segregation by sex, science and engineering stand out as the stubborn exceptions to the general trend of significant progress toward gender equity.

Why is it that science and engineering occupations seem less responsive to the social forces that have successfully effected progress toward gender equality in other professional occupations? Is science the "final frontier" for occupational gender equality? Is it simply that more time is needed, or is science so unique that it defies the trend toward gender equality witnessed in other professional occupations? Our task in this book is to shed new light on the question of why women continue to be underrepresented in science by systematically examining and documenting the career processes and outcomes of women scientists and comparing their experiences to those of men. We accomplish this task by departing from the common "science pipeline" approach and adopting the life course approach to the study of science careers. By calling attention to the multidimensionality of human lives as well as to the multiple levels of influences that combine to affect career trajectories and outcomes, the life course perspective provides a new conceptual and methodological framework that guides our study of women in science.

At the risk of oversimplification, let us borrow the economist's language and classify factors that affect women's underrepresentation in science into "demand" factors and "supply" factors. Demand factors are characteristics associated with jobs and employers that discourage women from entering an occupation. Such characteristics could include overt or covert gender discrimination. Supply factors are individual-level characteristics that deter women from pursuing science careers and/or encourage their pursuit of another type of career. Supply factors include educational background, career preferences, and family status. This dichotomy is illustrative and far too simplistic, for there can be important interactions between demand and supply influences. For example, women may consciously avoid attaining the prerequisite scientific education if they anticipate future difficulties, such as gender discrimination, in pursuing science careers.

We contend that the relatively slow movement of women into science and engineering occupations is caused by the forces of both supply and demand. This is significant because much of the blame for women's underrepresentation in science has been attributed to the practice of science per se or to the actions of male scientists. As we will demonstrate throughout the book, the inadequate supply of interested and qualified women has been as much, if not more, of a hindrance to the feminization of science as has the influence of demand factors. In other words, in our view, it would be naïve to presume that science and engineering occupations are closed to women simply through discriminatory practices and structural barriers. Despite a trend toward convergence, it remains true that women trail men in their desire and persistence in pursuing scientific careers. Among those who do pursue science, however, a significant portion of women achieves a level of success on par with their male colleagues.

The persistent enigma of women's underrepresentation in science has fueled a vast body of research. Scholars have examined a variety of questions about women's participation in, exclusion from, and contributions to the fields of science and engineering. Despite the significant breadth and depth of this research, much of it suffers from conceptual and methodological limitations that restrict the significance and usefulness of its findings. As a consequence, we have only limited knowledge of the processes that produce the gender differentials in science participation and attainment. In this book, we address the gap by presenting the first systematic examination of gender differences in the science career trajectory throughout the life course: from middle school through the career years. Our research explores both the early life course processes of selection into and out of the science educational track and the stratifying influences that operate after entry into the science labor market. The results of this study contribute to the literature on women in science, as well as to the broader literature on labor force gender inequality.

Our study encompasses the entirety of a career trajectory and provides a comprehensive, updated, and systematic empirical account of where and how women fall behind men in pursuing scientific careers and in practicing science. However simple this task may seem, it is in fact extremely difficult to execute. The project has required us to analyze seventeen large, nationally representative datasets (including the 1960-1990 U.S. Censuses) at the individual level. Any empirical social scientist would surely appreciate the sheer scale of the project. Throughout the project, our philosophy has been to learn as much as we can about the empirical world and to impart this valuable knowledge to our readers, each of whom may come to our book with different orientations and may act in different ways on the information conveyed in these pages. In this respect, our approach is decidedly demographic, for it values empirical knowledge for its own sake while allowing for alternative interpretations. In our view, the large literature on women in science has been plagued by broad discussions that pay insufficient attention to factual information. We hope to balance this tendency of the literature on women in science by providing the necessary empirical foundation for the discussions to move forward.

The design of our study addresses many of the conceptual and methodological limitations of previous research on women in science. We depart from the common "pipeline approach" to the study of science careers and instead draw on the life course perspective to frame our research and to guide the organization of this book. The result is a study of career processes and outcomes that emphasizes the human ecology of career development. We examine, for example, the causal influence of prior experiences on later career outcomes and the interactions of the multiple domains of an individual's life such as career and the family. We accomplish this by relying extensively on the statistical analysis of longitudinal data and, in places where true longitudinal data are lacking, "synthetic" cohorts constructed by piecing together information from different sources pertaining to different periods of the life course.

Although limited to science and engineering professions, our study is a concrete example of how to understand gender differences in career processes and outcomes in the general labor force. The innovative and expansive design of our study can be easily adapted for the study of gender inequities in other professions. The empirical findings reported in this book have important implications for, if not direct generalizability to, gender differences in other high status occupations. It has long been recognized by researchers and felt by individual women that gender differences in career trajectories are intimately linked to gender differences in the timing of events, prioritizing of roles, and social relations across the life course (Hochschild 1994; Maccoby 1995). Our research analyzes a career trajectory in its entirety in an attempt to illuminate the life course processes that at some times facilitate and at other times inhibit the career development of women relative to men.

Why Study Women in Science?

There are two primary social problems motivating the study of women in science. The first is the persistence of gender inequality in the labor force and the role of the gender segregation of occupations (including scientific and engineering occupations) in its maintenance. Stratification researchers have long documented that science is one of the most prestigious segments of the labor force (Hodge, Siegel, and Rossi 1964), and that women's low rates of participation in science contribute to, in aggregate, the lower social status of women relative to men (Jacobs 1989; Reskin 1984; Reskin and Hartmann 1986). An increase in the representation of women in science would lessen occupational segregation and reduce the level of gender inequality in the labor force. Theoretically speaking, one may expect women's entry into science to be relatively easy, given that science is lauded as relying on universalistic criteria for the evaluation and commendation of its members and their work (Cole 1979; Cole and Cole 1973; Merton 1973). If universalism is truly a norm in science, the problem of sex discrimination should be absent or at least less prevalent there than in other occupations (Xie 1989). We should, therefore, expect women to prefer science as a channel for upward social mobility. In reality, women's relatively slow movement into science suggests that career processes in science are influenced by multi-faceted social forces far more complicated than those dictated by the simplistic universalism principle.

Concern in the U.S. about the supply of science labor provides the second motivation for studying women in science. To a large extent, the booming U.S. economy of the past half century has been driven by advances in science and technology. Keeping pace with these advances and ensuring their continuation requires the maintenance of a scientific labor force of adequate size and quality. Concern is frequently raised about a possible shortage of scientists in the U.S. (Atkinson 1990).[1] Recruiting more women (along with underrepresented minorities) into science is often proposed as an effective way to expand the talent pool and therefore to increase the supply of scientists (National Science Foundation 1999). An implicit assumption underlying this policy approach is that the distribution of talent for scientific work is unrelated to gender, although recruitment into science traditionally has been. All available evidence suggests that innate differences in scientific ability between men and women are non-existent or very small (Maccoby and Jacklin 1974). Most scholars attribute women's lower representation and lesser success in science to social structural, social psychological, and family-related barriers (Sonnert 1999; Valian 1999; Zuckerman 1991). Thus the argument is: If we could remove or lower these barriers, we would expand the talent pool and increase the supply of well-trained scientists and engineers.

As social scientists and private citizens, we see gender-based inequality in the labor force as a social problem requiring remedy. We would like to see women's talent tapped more fully for science, both for the personal satisfaction of individual talented women as well as for the benefit of society as a whole. Although our research is motivated by these personal preconceptions, their influence stops there. We try to be objective and "value-free" (in so far as it is possible) in our empirical research. We conducted the statistical analyses for this study with great care, and we make every effort to report the empirical results in a balanced manner. In many places, we resist the temptation to impute our subjective interpretation into data, preferring instead to draw empirically grounded conclusions that are informed by the life course perspective. Since we wish to reach a wide audience, complicated statistical results are compressed and reported only in summary form in the main text. Detailed information about statistical results and data analysis is available in the appendices.

Previous Research on Women in Science: The Legacy of the Pipeline

The past two decades have witnessed a proliferation of literature on women in science. Indeed, a casual search in a major research library would easily yield more than a dozen book titles specifically addressing women in science, most of which were published after Cole's 1979 landmark book *Fair Science: Women in the Scientific Community*. (See Davis et al 1996; Gornick 1990; Hanson 1996; Kahle 1985; Long 2001; McIlwee and Robinson 1992; Pattatucci 1998; Schiebinger 1999; Selby 1999; Sonnert 1995a, 1995b; Wasserman 2000; and Zuckerman, Cole, and Bruer 1991.) A search of scholarly journals in sociology, psychology, history, and women's studies would add a plethora of articles that address topics ranging from the existence and extent of gender differences in math-related brain functioning (e.g., Haier and Benbow 1995) to the influence of family characteristics on participation in math courses during high school (e.g., Eccles, Jacobs, and Harold 1990), and from gender differences in publication productivity (e.g., Xie and Shauman 1998) to the impact of feminization on salaries in specific academic disciplines (e.g., Bellas 1994). In addition to this body of academic research, government agencies have periodically published reports full of statistical information that measure the participation and progress of women in science education and occupations (e.g., Committee on Women in Science and Engineering 1991; National Science Foundation 1986, 1992, 1994, 1996, 1999, 2000).

The sheer size of this literature indicates that there is a great deal of interest in the subject of women in science in both the scholarly community and the public at large. This large and growing body of research has illuminated many aspects of women's careers in science and the ways in which the careers of men and women scientists differ. Some scholarship also argues that women's participation in science itself enriches scientific thinking by bringing more subjectivity, more empathy, more reliance on intuition, more holism, and perhaps more passion to science than has been practiced by men.[2] We will not address the philosophical issues of women's increased involvement in science in this book. Instead, our research engages the literature that focuses on gender differences in science career trajectories and their causes. Due to conceptual drawbacks and associated methodological limitations, the extant research on the careers of women in science has failed to answer many old questions, and indeed it has raised new ones. In the sections that follow, we briefly discuss some of the inadequacies of the literature before describing our approach to the study of women in science.

Limitations of the "Science Pipeline" Conceptualization

The major conceptual limitation of the literature on women in science is the predominance of the "pipeline" model. According to this model, the process of becoming a scientist can be conceptualized as a pipeline, called the "science pipeline." The image of the science pipeline is used to illustrate a structured set of educational and employment stages that comprise a science career. The pipeline typically refers to the sequence of college-track math and science courses in middle and high school, followed by science concentration in undergraduate college, science graduate study, and/or employment in a science occupation. The pipeline conceptualization posits a straight and narrow connection between education and occupation and prompts the conscious consideration of the career as a developmental process encompassing both educational and occupational outcomes. A common conclusion drawn from this conceptual perspective is that the underrepresentation of women in science is attributable to women's relatively higher rates of attrition from the science pipeline. Hence, the implication of the pipeline model research is that, to increase women's representation in science, policies must be devised to "block" the leakage at those points where the pipeline loses more women than men.

Berryman's (1983) introduction of the pipeline model to empirical analysis was a major innovation that facilitated the examination of gender differences across the entire process of becoming a scientist, instead of focusing on single career stages in abstraction. For example, Berryman analyzed gender and racial differences in the science career trajectory by compiling detailed statistics by gender and race that pertain to attainment of sequential levels of science education. From this analysis, Berryman (1983, p.5) concluded that "For women, the losses are concentrated at the end of the pipeline: at the Ph.D. level." We note that Berryman's study was based on cross-sectional data, as are most other studies that are based on the pipeline conceptualization.

Since Berryman's study, the pipeline model has become so widespread that it is commonly accepted as the dominant, if not the standard, conceptual framework within which to organize studies of the science educational and career trajectory. For instance, an overview article published in *Science* was provocatively entitled "The Pipeline is Leaking Women All the Way Along" (Alper 1993), and two recent reviews (Schiebinger 1999; Sonnert 1999) also include "pipeline" as part of a major heading. With the exception of Hanson (1996, p.4), the pervasive use of the pipeline model has gone unquestioned. This is unfortunate, since the pipeline model, while a useful conceptual framework, has limited empirical research by restricting the kinds of questions that are asked by researchers.

To argue that the pipeline framework should not dictate the research agenda, we point to three major limitations. First, the pipeline model does not capture the complexity of the educational and career processes of becoming a scientist. It refers to a unidirectional, orderly, and rigid series of stages, and it equates noncompliance with the normative career trajectory to "leaking" or "dropping out" of the pipeline. This characterization is reified by research that narrowly focuses on gender differences in exits from science/engineering (denoted as S/E) at the expense of not studying the actual, often complicated, pathways followed by men and women in the pursuit of science careers.[3]

Second, in the pipeline framework, persistence across different stages of the educational and career trajectory is assumed to represent progress along the science pipeline. In other words, the pipeline model is a developmental framework in which the successful completion of all stages within an ideal-typical time schedule means a positive outcome. Nonparticipation at any stage is equated with dropping out of the pipeline, and movement back into the pipeline after dropout is assumed to be structurally improbable or impossible. Thus, the narrowness of the pipeline conceptualization has precluded the consideration of alternative educational and career trajectories, neglecting, for example, a thorough examination of the possibility and implications of "late" entry into S/E and whether the existence or nonexistence of such a career path influences gender differences in participation.

Third, other life course events, such as family formation, that coincide and interact with the science career trajectory are absent from the pipeline conceptualization. By not situating the science career in the context of other life course events, pipeline researchers have implicitly assumed that the pipeline is independent of the timing and character of other life course events. As a result, past research is mostly individual-centered and overlooks the role of the family, when gender differences in family expectations and the demands of familial roles may have a significant impact on the timing and sequencing of women's science careers.

Methodological Limitations

Associated with the conceptual drawbacks of the extant literature are some common methodological limitations. First, the current literature on women in science pays inadequate attention to the complexity of measuring career processes and outcomes. Since the processes and outcomes that comprise a career are multi-faceted, it is a gross oversimplification to focus on just one or two aspects of the career in investigating women's experiences in science. To be sure, a few studies have been very thorough (e.g., Ahern and Scott 1981; Cole 1979; Long 2001; Sonnert 1995a). These studies, however, are each based on a single data source and are consequently narrow in scope. The career outcomes typically considered are productivity, rank, and salary. In our study, we attempt to address gender differences in career processes and outcomes using a variety of measures within a single broad conceptual framework. In doing so, we update old results and produce new findings.

A second methodological limitation is the widespread use of select, nonrepresentative samples.[4] For example, Sonnert's (1995a, 1995b) two influential books were based on a database about former postdoctoral fellows who received prestigious fellowships from the National Science Foundation or the National Research Council. Other examples include Etzkowitz et al.'s (1994) study, which collected data in four departments at a single research university, and Wasserman's (2000) project, which is based on interviews with women who have been elected to the National Academy of Sciences. These focused studies do provide good insight into why women trail men in science careers. In some ways, they may be better than large, nationally representative databases at enabling a deeper examination of the institutional, cultural, and structural barriers that women scientists may face. At the same time, however, studies based on nonrepresentative samples have the irremediable shortcoming of producing results that are not generalizable to the reference population of women scientists.

The widespread use of nonrepresentative, nonrandom samples in studies of women in science underscores the difficulty of conducting a formal statistical analysis of women scientists: there are relatively few of them in the population. Random sampling of the national population is therefore an extremely inefficient way to collect the necessary data. Researchers therefore often collect data from a well-defined subpopulation, such as recipients of doctoral degrees in science (e.g., Cole 1979; Long, Allison, and McGinnis 1993). While sampling from a subpopulation is effective—indeed it is an approach that is also adopted to a large extent in this book—it is important to realize that this approach also carries some methodological risks in light of the dynamic nature of career processes and outcomes.

One methodological problem with sampling from a subpopulation is the arbitrariness inherent in defining a static subpopulation in a dynamic process. Whether or not someone happens to be included in a subpopulation at a given time is undoubtedly affected by chance, and this approach allows chance to play too large a role in defining the sampling frame. When we draw a sample from all doctoral recipients in a specific year, we exclude not only individuals who obtained their degrees in the adjacent years but also those who have not obtained (and will not obtain) a doctoral degree but who are making important scientific contributions. Similarly, when we study academic scientists, we exclude scientists who work in the government and industrial sectors, scientists who have recently moved out of academia, and scientists who are temporarily out of the labor force. If the element of chance is entirely random in its effects on who is included in a particular subpopulation at a particular time, the problem is only a nuisance but does not introduce a serious bias to statistical results. Unfortunately, individuals with certain characteristics may be differentially affected by this uncertainty. We discuss this problem in greater detail below but wish to emphasize here that sampling from a subpopulation is not a simple matter, since determining at what point of the career process a sample should be drawn is a consequential decision.

A related methodological risk is the potential problem of selectivity due to what is commonly called “left censoring”: only individuals who have successfully progressed to a certain career stage are included in a sampling frame, and these individuals may differ systematically from those who previously dropped out for whatever reason. For example, when very few women worked in science, it was reasonable to assume that these women may have possessed characteristics that distinguished them from the vast majority of women who did not work in science. Since women have increased their representation in science, the degree of distinction between women working in science and those working in nonscience may have decreased. In this sense, we may want to say that the population of women scientists has become less selective with the increased representation of women in science. Theoretically speaking at least, selectivity could seriously bias statistical results based on observational data and thus render the results unreliable. Unfortunately, selectivity bias is always a potential threat to any study of scientists, because scientists constitute only a very small (albeit increasing) proportion of the labor force.

We do not claim to solve these methodological problems in this book. In the broad scheme, these problems are not solvable with observational data. Our solution is to be comprehensive and pragmatic: we analyze samples that are drawn at different stages of the life course starting from a very early age. In so doing, we benefit from having rich information about a certain segment of the educational and career trajectory from a single sample. We then juxtapose statistical results from different samples to gain a fuller understanding of the issues. Since our study covers the career trajectory from middle school onward, we gain knowledge about gender differences in earlier career stages before proceeding to analyze gender differences in career outcomes at later stages. Thus, although each separate analysis in our book may suffer from the common problems of left-censoring and selectivity bias, the project itself addresses the problem by using multiple data sources to cover the entire career process.

Another common methodological problem in the literature on women in science is the over-reliance on cross-sectional data, despite the fact that most hypotheses in the literature are about dynamic processes that require the analysis of longitudinal data. A prime example is Berryman's (1983) study, which provides excellent snapshots of women's representation in the science pipeline at different educational stages. As we argued earlier, the science career should be conceptualized as a dynamic process. Given the dynamic nature of the subject matter, longitudinal data should in principle always be preferred over cross-sectional data, although we recognize the limited availability of longitudinal data suitable for the study of women in science.

Finally, much of the earlier research on women in science has relied on simple, descriptive statistics (for example, Ahern and Scott 1981, Berryman 1983, Davis et al. 1996, and the special reports by the Committee on Women in Science and Engineering 1991 and the National Science Foundation 1999). While studies reporting descriptive statistics are informative, they can be misleading due to the possibility that observed gender differences in career outcomes may be confounded by other factors that are related to both gender and career outcomes. In a later section of this chapter, we illustrate the importance of multivariate analysis for controlling factors that mediate the observed association between gender and a career outcome. Throughout the book, we attempt to examine the net effect of gender in a multivariate framework and to compare multivariate results to descriptive results.

Life Course Perspective

The life course perspective frames our examination of the career processes and outcomes of women in science and organizes the presentation of this book. The life course perspective shares with the pipeline perspective an attention to the sequencing and interdependence of educational and occupational events, but the similarity is a superficial one. Whereas the pipeline perspective assumes the science career to be an exceptionally rigid structure, the life course perspective allows for a more multidimensional and nuanced understanding of career processes and outcomes. In a nutshell, the life course perspective posits that the occurrence of significant events and transitions in an individual's life is age-dependent, interrelated, and contingent on (but not determined by) earlier experiences and societal forces.

Researchers have long recognized that inequality in the labor force is the manifestation of a process of stratification that occurs throughout the life course (e.g., Merton's [1973] theory of cumulative advantage). In particular, the persistent effect of educational attainment on subsequent occupational attainments has been a core subject of traditional social stratification research. However, empirical analyses of specific career lines have tended to separate the educational career from the labor force career and thus have failed to connect them in a life course conceptualization. To understand gender differences in science careers, career lines must be recognized as "lifetime trajectories which result from stratification processes occurring in both educational and labor force settings" (Kerckhoff 1996, p.38). In this study, we conceptualize the S/E career trajectory as a life course process that begins with S/E education and extends to S/E labor force participation. Our investigation is aimed at understanding and gauging gender disparities during different periods of the life course and examining how early gender differences establish and reinforce gender inequalities later in life.

Conceptualizing the S/E career trajectory as a life course process highlights the need to model explicitly the dynamic processes that lead to the attainment of educational credentials in science, placement in scientific occupations, and the experience of successive career outcomes. The emphasis on the dynamic processes of the life course has guided our research in four ways.

First, the life course perspective recognizes the interactive effects of institutional-level and individual-level influences on the course of career trajectories. That is, "which individuals follow which trajectories is determined by the intersection of individual and institutional actions" (Kerckhoff 1996, p.38). We argue that gender inequalities in S/E careers are produced by the interaction of structural allocation and self-selection processes. We use the word "self-selection" because we assume that humans lead purposive lives: individuals set both short-term and long-term goals about the roles they intend to fill, and they act to achieve those goals (Clausen 1986). However, it is also important to recognize that the range of considered options is largely constrained by the social and cultural norms that are reflected in the social structure and reinforced by the significant actors in one's life (Clausen 1986; MacLeod 1987; Sewell, Haller and Portes 1969; Xie and Shauman 1997). Individuals perceive their options through their unique "matrix of social characteristics" (Xie 1989, p.4). The objective of this research is to discover how certain configurations of factors lead some individuals, but not others, to believe that the scientist role is desirable and attainable, to maintain and act on that belief through an extended period of the life course, and why this sorting of individuals into S/E and non-S/E pools so neatly divides along gender lines.

Second, the life course perspective attempts to model the complexity of human lives by recognizing that the life course is composed of multiple trajectories in the domains of education, family, and work. Progress in a career trajectory is always accompanied, and may be influenced, by developments in other life course trajectories (Elder 1977; O'Rand 1996; O'Rand and Krecker 1990). Thus, educational and occupational achievements and choices are affected by developments in the family domain as well as by individual and structural influences that are associated directly with education and occupation. Events in one domain of the life course can affect the course of trajectories in other domains through their effects on one's available time, interest, energy, and material resources.

Third, the life course perspective recognizes the existence of "career lines," or relatively structured and frequently traveled pathways through multiple transitions to particular life course outcomes, but it also emphasizes the individual-level variation in career tracks. It is assumed that there are normative pathways linking social-structural origins to career destinations. However, the life course perspective also directs us to recognize that there are many other viable career tracks. In fact, as we show in this book, for women the less-traveled pathways may be important routes to science. The research agenda set by the life course perspective aims to identify the "most frequently traveled pathways" to S/E careers (Kerckhoff 1996, p.38) and to understand the systematic variation at the individual level in these paths.

Fourth, the life course perspective points to the cumulative nature of the life course by which small differences at particular points in time can deflect trajectories and subsequently generate large differences in career outcomes. The idea that the stratification of career outcomes is the culmination of small differences at earlier points in the life course was introduced by Robert Merton's (1973) cumulative advantage hypothesis (also see Cole and Singer 1991). Kerckhoff's (1993) work documents that modest influences can accumulate into substantial "deflections" over the life course. Deflections are produced and reinforced at the institutional level through structural allocation, and at the individual level through socialization, actualization, and self-legitimization. Accurately identifying the causes of gender differences in career trajectories requires a holistic research design that examines a significant span of the life course and multiple levels of social influences.

In short, a key strength of the life course perspective is its recognition of the multidimensionality of individuals' lives. It encourages social researchers to recognize the multiple roles that characterize individual lives, the layers of forces that impact an individual's life course, and the succession of events and outcomes that comprise a "career." For our study, we outline a schematic model that specifies the dynamics of a particular set of social determinants that operate over the course of someone's life and affect his or her educational and occupational outcomes. For convenience, we group these social determinants by the level of aggregation into individual influences, familial influences, and broader social influences. We propose that the impact of these determinants gradually shifts in measurable and predictable ways as individuals move through the life course (Maccoby 1995). We also propose a set of career processes and outcomes that are appropriate to particular periods of the life course. While the set of career processes and outcomes we include in this study falls far short of complete coverage, the events we examine are theoretically important "points of deflection" in the S/E career. This schematic model serves as a heuristic guide for our empirical study rather than as the theory to be supported or rejected. In the following subsections, we discuss the three groups of social determinants in greater detail and present the sequence of career processes and outcomes that organize our life course analysis of women's careers in science.

Individual Influences

Early in the life course, certain individual characteristics such as intelligence or career ambition should matter a great deal for educational outcomes. For example, influential throughout primary and secondary school is an individual's demonstrated aptitude, that is, her/his ability to learn and to perform well on classroom work and achievement tests. Later in the life course, academic performance remains influential, but the character of its influence changes. As an individual progresses along the life course from high school to college and the labor force, his/her educational and occupational trajectory is influenced less by aptitude and more by mastery of the skill to perform specific tasks. As new scientists enter the labor force, their career prospects are judged more by the credentials they have attained in formal education rather than by any measure of cognitive ability. Needless to say, attainment of educational credentials depends on cognitive ability. However, the causal effects of cognitive ability on educational credentials are by no means deterministic. What we argue here is that most of the effects of cognitive ability on science careers are indirect, mediated by formal education. Furthermore, some new individual characteristics emerge as influential as individuals begin to participate in a labor market. For example, characteristics such as the time between bachelor's degree and doctorate and access to research funding are factors that become influential for career paths only at or after entry into the labor force.

The expected shifts in individual influences throughout the life course inform our analysis and are reflected in the selection of explanatory and outcome variables for each chapter of the book. In the early chapters we explicitly examine gender difference in measured achievement in math and science and then go on to analyze the influence of the math and science achievement on other outcomes such as the formation of S/E career aspirations and the choice of an S/E college major. These chapters also include measures of affective characteristics such as comfort with math and science and attitudes about the perceived future usefulness of math achievement. The later chapters of the book address gender differences in labor force experiences and outcomes. The analyses of these later chapters include explanatory variables that are expected to be most relevant for individuals participating in the S/E labor market. For example, the analysis of gender differences in publication productivity in Chapter 9 examines the influence of individual characteristics such as time to degree, access to research funding and assistance, and time spent on classroom teaching.

Familial Influences

Changes also occur in the nature of the familial influences. A general trend is that the locus of familial influences shifts from the family of origin (i.e., the natal family) to the family of formation (i.e., the primary family) as the individual progresses from childhood to adulthood (Maccoby 1995). Early in the life course, one's parents are the primary socializing agents. Youth learn the social rules and norms and are oriented to their world and taught their place in it by their parents and other members of the natal family. Parents make most decisions for their children, and the few decisions that children make on their own are strongly influenced by the beliefs of their parents. As the individual matures, his/her parents modify their treatment (Maccoby and Jacklin 1974), slowly yielding their authority, and gradually reducing their influence. Eventually, individuals find themselves making their own decisions. As this transition occurs, the direct influence of parents wanes, and life decisions are increasingly influenced by an individual's conceptions of his/her own current and future social roles. Prominent among these conceptions is the orientation to the gender roles society prescribes, especially the roles associated with marriage and family. As individuals later marry and begin to have children, their life decisions become increasingly influenced by the familial roles they assume and the needs of the other members of the primary family.

Before individuals marry and start a family of their own, the influence of familial concerns is manifest in the influence of sex-role socialization on a young person's perceptions of the current and future social costs of pursuing a career in science. The anticipated conflict between a demanding career and childrearing may be a salient issue for girls early in the life course, long before they actually experience childbearing and childrearing. For young girls who plan to combine work and family roles, the heavy human capital investments and anticipated time demands of S/E occupations may squelch their interest in an S/E career (Sandberg et al. 1987; Wolfe and Betz 1981). Later in the S/E career trajectory, familial influence is manifest in the tension between the actual familial life and the commitment to a career that requires heavy investment in education, training, and role performance. Reflecting this shift in familial influences, variables measuring the expected timing of family formation and the relative importance of familial roles are central to the analyses of the early life course presented in Chapters 2, 3, and 4.

Beginning with the analysis of bachelor's degree attainment in Chapter 4, we introduce marital and parental statuses as explanatory factors and include them in later chapters whenever data permit.

Social Influences

Although parents and the family lay the foundation for career development, individuals are also exposed to influences through social institutions that operate outside of the family. When individuals are young, they are mostly affected by the school system. As schooling ends and work begins, the locus of social influences shifts from schools to the labor force.

Schools exert their influence in a variety of ways: the availability of courses, sporting facilities, and extracurricular activities; the quality of teaching, teachers' expectations and guidance; the availability and orientation of guidance counselors; and the influence of peers. Although in principle it is possible for schools to initiate social change by treating girls and boys equally, in practice schools often serve to reinforce traditional gender stereotypes. This is true because schools often reflect the social structure of the larger society. For example, very few science and math teachers are women, and both male and female teachers report having higher expectations for girls than boys in language-related tasks and higher expectations for boys than girls in math and science activities (Reid and Paludi 1993). Whereas the influence of the educational structure is limited to the period of direct contact, the influence of the labor force structure may begin even before individuals officially enter it. Students are aware of the structure of the labor force and use that information to plan their further education and careers, and to assess the appropriateness of their aspirations (Xie and Shauman 1997). Students' career aspirations reflect their perceptions of the careers that are appropriate for and attainable by them. Therefore, gender differences in aspirations reveal the degree to which young men and women perceive different opportunity structures for themselves (Laws 1976; Marini and Brinton 1984). The fact that young women are much less likely than young men to aspire to an S/E career indicates that a significant proportion of young women do not perceive S/E occupations to be within their realm of possibility, and this perception is reinforced by the structure of the labor force.

Once individuals make the transition into the labor force, their actions, access to resources, and achievement-related goals are influenced by their work setting and their coworkers. Different work environments lead to different and unequal career outcomes. For example, women academic scientists are more likely than men to be found in teaching colleges rather than in research universities (Fox 1995, p.212; Long 2001; Long and Fox 1995), and this helps explain gender differences in publication productivity (see Chapter 9).

Processes and Outcomes in S/E Careers

Just as the life course perspective draws our attention to the multiple levels of influences that operate throughout the course of a career, it also calls for an examination of career processes and outcomes that are appropriate to particular stages of the career trajectory. A distinct feature of this study is our attention to the multi-faceted nature of career processes and outcomes. This feature reflects our belief that gender differences in S/E career processes and outcomes accumulate over the life course and manifest themselves at different points in the life course. We cannot, however, include everything that is relevant to this subject. Some topics are excluded from the book because they are relatively less important than those that we choose to study. Other topics are omitted because we do not have satisfactory data with which to address them. The aspects of career processes and outcomes that we study are:

(1) Academic achievement in science and mathematics in pre-college years.

(2) The expectation of enrolling in an S/E major in college.

(3) The likelihood of obtaining an S/E degree at the bachelor's level.

(4) Career outcomes following the completion of a bachelor's degree in S/E.

(5) Career outcomes following the completion of a master's degree in S/E.

(6) The demographic and labor force characteristics of scientists/engineers.

(7) Geographic mobility of doctoral scientists/engineers.

(8) Research productivity among academic scientists.

(9) Immigrant women scientists.

Clearly, the above list falls far short of exhausting all the possible topics that can be considered as career processes and outcomes. Another caveat is that, although we earlier discussed the influence of social determinants in connection with the life course approach, our analysis of educational outcomes in S/E is mainly concerned with individual-level and family-level influences. We do not examine school-level and other broad social influences in our study. Our decision to limit the scope of this research prevents us from directly testing the hypothesis that gender differences in S/E education are affected by such school-level social influences as teacher behavior, instructional style, and classroom atmosphere. Rather, we can only indirectly infer the influence of the school from the effects of individual attributes and family characteristics. Our decision is based on the limitations of our data as well as our desire to accomplish a thorough investigation of the individual and familial influences. We leave the role of the school-level influences to be explored by other researchers.

Methodological Issues

Implementing this study of the career processes and outcomes of women in science required many decisions about the research design, concept definitions, and methodological issues. Essentially, the book consists of nine separate analyses targeted at different topics. While a specific analysis often required unique methodological decisions, some common methodological issues permeate the entire project. In this section we address the methodological concerns and elements of the research design that are common to all of the component analyses of the study.

Synthetic Cohorts

From the previous discussion, it should be clear that the life course framework calls for longitudinal data covering individuals' educational and career histories. In reality, however, we have access only to cross-sectional or longitudinal data of limited duration. To overcome this difficulty, we devise a demographic approach to studying the developmental process of becoming a scientist/engineer that follows a synthetic cohort through the formative years of career development. The approach is dynamic rather than static in the sense that it traces changes over the different career stages of a cohort.

In a classic article, Ryder (1965, p.845) defined a cohort "as the aggregate of individuals (within some population definition) who experienced the same event within the same time interval." For example, all individuals born at the same time (say within a given calendar year) make up a birth cohort. Similarly, events such as marriage and school entry define marriage and school cohorts. Ideally, we would like to observe all career changes of a real cohort for its entire history, from childhood to retirement. This would allow us to accurately model the life-course career process of the cohort. Such longitudinal designs, however, are unrealistic in practice not only because they are too expensive and would yield extremely small samples of practicing scientists, but also because they take a lifetime to complete and thus cannot yield even tentative answers to important questions the society currently faces.

One common solution to this dilemma, often adopted by demographers in studies of fertility and mortality, is to construct age-specific vital rates from a cross-section and then assume them to be experienced by a hypothetical cohort. For instance, the total fertility rate (TFR) is the expected total number of children a woman would have if she followed the entire age-specific fertility schedule of a given period, and life expectancy is the expected total number of years a newborn child would live if she/he were subject to the age-specific mortality schedule of a given period. Berryman's study (1983) is an application of this approach. As we discussed above, however, this approach does not allow researchers to uncover dynamic processes underlying the cross-sectional data. For example, Berryman was unable to examine the changes in enrollment status and field of study and their variations across gender and race, even though she clearly realized the importance of such transitions.

Limited longitudinal studies, a middle ground between purely cross-sectional designs and ideal longitudinal designs, have gained more popularity and acceptance in recent years. By “limited longitudinal studies” we mean that a group of subjects is followed only for a limited duration. Examples are the National Longitudinal Study of the Class of 1972 (NLS-72), Longitudinal Study of American Youth (LSAY), High School and Beyond (HS&B), National Educational Longitudinal Survey (NELS), and the Survey of 1982-1989 Natural and Social Scientists and Engineers (SSE). We use these limited longitudinal studies for our research and describe them in Appendix A.

Limited longitudinal studies could be cohort-based, such as NLS-72, LSAY, HS&B, and NELS, or population-based, such as SSE. While the sampling frames of NLS-72, LSAY, HS&B, and NELS were school cohorts, the sampling frame of SSE was the population of scientists identified by the 1980 U.S. Census. Because there are currently many large, nationally representative, cohort-based limited longitudinal studies available, we piece together the experiences of different cohorts to form a synthetic cohort. Here we define a synthetic cohort as a hypothetical cohort whose life history is constructed from different real cohorts in a supplementary manner. Although the synthetic cohort is not real, segments of the cohort's experiences are real. A major advantage of this approach is that it allows us to study the “social dynamics” of individuals’ transitions into and out of different educational and career states (Tuma and Hannan 1984).

Unfortunately, even limited longitudinal studies with reasonable sample sizes are difficult to come by. For some important phases of career development and some important career outcomes, we do not have access to even limited longitudinal studies. When this is the case, we resort to using cross-sectional data for snapshots of dynamic processes, and we sometimes rely on respondents’ recall of retrospective information. Despite the limitations of cross-sectional data, we believe that empirical findings from such data can be very informative. We therefore supplement longitudinal data with cross-sectional data in some chapters in this book. Together, the information from many different datasets contributes to a composite portrait of the career processes and outcomes of women in science. This composite portrait does not belong to any real person, nor any real cohort, but to a synthetic cohort. See Figure 1.1 for an illustration of the segments in the life course that we examine and the various datasets we use.

[Figure 1.1 about here.]

Synthesizing results pertaining to different cohorts as if they belong to a single cohort requires the assumption that the career process is mainly age-dependent rather than cohort- or period-dependent. This assumption is called "stationarity," referring to the robustness of age patterns across different cohorts and historical periods. Over fifty years ago Ginzberg and his associates (Ginzberg et al. 1951) observed that the career process is generally an age-dependent developmental process. As we show in this book, however, women's careers in science have changed rapidly in recent decades, and the assumption of stationarity required for interpreting the results as if they pertain to a synthetic cohort is likely to be violated in reality. Thus, the reader should treat the synthetic cohort approach discussed in this section as a heuristic device that aids in the interpretation of the empirical results from many datasets. Indeed, the empirical results reported in the book are meaningful in and by themselves, and the interpretation of them does not hinge on one's acceptance of the synthetic cohort approach.

Definition of Scientists/Engineers

Before we can proceed further, we need to define what we mean by scientists/engineers. At first glance, the task seems straightforward, for the terms "scientists" and "engineers" are commonly used. In practice, however, defining scientists/engineers is a difficult task facing all researchers who study the population (Citro and Kalton 1989).

There are at least three approaches to defining scientists/engineers. Let us call the three resulting definitions the substantive definition, the credential definition, and the behavioral definition. The substantive definition recognizes the existence of “invisible colleges” (Crane 1972)—the scientific community—in which scientists communicate with each other through formal and informal publication channels. The ultimate criterion for the substantive definition is one's contribution to a common body of scientific knowledge and thus one's involvement in scientific communication. The main problem with this approach is the difficulty of measuring contribution. In contrast, the credential definition is based on one's formal credentials, such as an S/E degree. This definition is also called the education-based or supply-based definition, as it defines the “supply” of scientists. The main problem with the credential definition is that not everyone with the requisite credentials is actually involved in scientific activities. Lastly, the behavioral definition takes the “demand” perspective and relies on information about activities that individuals actually perform in their jobs. We also call this definition occupation-based or demand-based since it defines a person as a scientist if he or she reports holding a scientific job. The populations of “scientists” identified by these three definitions do not coincide. A person may be labeled a scientist according to one definition but a non-scientist according to another. For example, it is possible to find a small number of people actually doing scientific work even though they neither hold formal degrees nor publish scientific papers. More often than not, however, an individual is identified as a scientist by all three definitions.

The National Science Foundation (e.g. NSF 1986, p.39) had a long-standing, eclectic practice that combined information pertaining to the three definitions.[5] Roughly speaking, it defined someone as a scientist/engineer if at least two of the following three criteria were met: (1) the person's highest degree is in science/engineering; (2) the person is employed in a scientific/engineering occupation; and (3) the person professionally identifies himself/herself as a scientist/engineer based on total education and work experience. There are a number of problems with this hybrid definition. The most serious of all is that it does not have a clear theoretical interpretation. Another is the difficulty of maintaining consistency across datasets. Finally, the third criterion, self-identification as a scientist, gives too much room to individual subjectivity.

We adopt two alternative definitions of scientists/engineers: the credential (education-based or supply-based) definition, and the behavioral (occupation-based or demand-based) definition, following the recommendation of Citro and Kalton (1989). In practice, the occupation-based definition specifies that the incumbents of S/E occupations are scientists/engineers. The education-based definition considers individuals with or working toward S/E educational degrees as scientists/engineers or potential scientists/engineers. In the first part of the book, we are primarily concerned with education in science and engineering and thus rely on the education-based definition. In the second part of the book, where the focus is on career outcomes of scientists/engineers, we shift to the occupation-based definition. Although this strategy leaves vague who we ultimately wish to define as scientists/engineers, it allows us to define the most appropriate population for the analysis of each specific aspect of career processes and outcomes that we study.

Measuring Gender Differences

Although we are mainly interested in the experiences and outcomes of women in science, we believe it is essential to include men in our analysis. While a few studies (e.g., Wasserman 2000) of women in science have been based on data from women alone, such studies are inherently limited in their ability to identify barriers unique to women's careers in science and engineering. As in most studies on women in science, we treat men as the natural reference group with whom to compare women. We maintain this focus on gender comparison throughout the statistical analyses reported in this book. Indeed, whenever feasible, we present only those statistics that contrast women and men. In this section, we discuss the pros and cons of several single-number measures summarizing gender differences.[6]

Correlation

Earlier work on women scientists (e.g., Blackburn, Behymer, and Hall 1978; Cole 1979) used the correlation coefficient as the primary measure of gender differences. To use this measure, the researcher typically codes sex as a dummy variable and then computes the Pearson correlation coefficient (*r*) between sex and a continuous outcome variable. The Pearson correlation involving a dummy variable for sex (*X*) and a continuous variable (*Y*) can be calculated easily from the sex-specific sample means of *Y*, the sex composition in the sample, and the standard deviation of *Y*. More specifically (Stuart and Ord 1991, p. 995),

$$r = \sqrt{pq}(\bar{Y}_{\mathrm{F}} - \bar{Y}_{\mathrm{M}}) / S_y, \tag{1.1}$$

where p and q denote the proportions of female scientists and male scientists in the sample, and $\bar{Y}_{\mathrm{F}}$ and $\bar{Y}_{\mathrm{M}}$ respectively represent the means for female and male scientists. Equation 1.1 reveals that the correlation between *X* and *Y* is <u>not</u> invariant with respect to the sex composition in the sample. Specifically, the correlation reaches the maximum possible value when the sex composition is balanced at 50 percent female and 50 percent male and declines when the sex composition is unbalanced. This is so even though the essential information (i.e., sex-specific means) in the data remains unchanged. For this reason, correlation generally is not a desirable statistic for measuring gender differences, especially in S/E where gender composition has been unequal but changing rapidly in recent decades.

Means Ratio

Since conditional means by sex convey the most essential information on sex differences, we can measure sex differences simply by taking the ratio between sex-specific means:

$$R = \bar{Y}_F / \bar{Y}_M \ . \tag{1.2}$$

The sex ratio of means (*R*) is invariant to changes in the sex composition. For most research questions, this invariance is desirable, and thus *R* is preferable to r. Indeed, this is the measure used in the studies of the gender differences in publication productivity by Cole and Zuckerman (1984) and Long (1992). Easy to compute and interpret, the ratio expressed by equation 1.2 has been the standard measure used in the labor force literature studying sex differences in earnings (e.g., Bianchi and Spain 1986). As will be shown later, this measure can also be computed easily from multivariate regression models with the logarithm of Y as the dependent variable.

Probability Ratio

The outcome variable (Y) itself may also be a dummy variable, taking the value of 1 or 0. An example is whether or not a person completes a bachelor's degree in S/E. When this is the case, equation 1.2 changes to probability ratio:

$$PR = \mathrm{P}(Y = 1|\mathrm{F})/\mathrm{P}(Y = 1|\mathrm{M}), \tag{1.3}$$

where $\mathrm{P}(Y = 1|\mathrm{F})$ is the probability among females, and $\mathrm{P}(Y = 1|\mathrm{M})$ is the probability among males. The probability ratio is also called the relative risk (Powers and Xie 2000, pp.94-95). Although the measure can be straightforwardly defined as an extension of the means ratio, it has the drawback of asymmetry. We know that the 0-1 coding for a dummy variable is arbitrary. Thus, an ideal measure of sex differences should not vary however we code the outcome variable Y. Unfortunately, the probability ratio is not invariant to which category of the dichotomous outcome variable is coded to 1 (versus 0) (Powers and Xie 2000, p.95). This drawback makes the measure a poor candidate for summarizing gender differences.

Odds Ratio

For dichotomous outcome variables, we resort to the odds ratio (OR) to measure gender differences. Simply put, odds ratio is the ratio of odds instead of probabilities, with odds (O) defined as

$$O = \mathrm{P}(Y = 1)/\mathrm{P}(Y = 0). \tag{1.4}$$

The odds ratio measuring gender differences in the likelihood of $Y = 1$ versus the likelihood $Y = 0$ is:

$$OR = [P(Y=1|F)/P(Y=0|F)]/[P(Y=1|M)/P(Y=0|M)]. \quad (1.5)$$

Unlike a probability ratio, an odds ratio is symmetric around the 0.5 probability point and invariant to changes in the coding of the outcome variable: the switch of the coding of $Y = 1$ and $Y = 0$ simply means the inversion of the odds ratio. When $P(Y = 1)$ is very small, $P(Y = 0)$ is close to 1, and a probability ratio and an odds ratio are thus about the same.

What is particularly attractive about odds ratios is that they can be retrieved easily from multivariate logit regression models. This is true because the dependent variable in a logit regression model is the logarithm of odds. Hence, we can compare observed odds ratios to adjusted odds ratios after controlling for covariates. In Chapter 7, we will also explore the relationship between an odds ratio as defined above and a representation ratio.

Explaining Gender Differences: The Multivariate Approach

Given our primary interest in comparing women and men, can we simply present statistics (such as R or OR) that pertain to gender differences on average? That is, can we conduct bivariate analyses showing relationships between gender and relevant outcome variables? The answer is no, although for the sake of parsimony we would like to.

The danger in drawing conclusions based on simple female-male comparisons is that such comparisons may be confounded by other relevant factors. To illustrate this point, consider the example of gender differences in earnings. Women earn less than men in part because women are more likely to work part time, and part-time workers earn less than full-time workers on average. A simple comparison of the average earnings of women and men thus would be misleading. In general, to tease out confounding factors that mediate between gender on one hand and educational and career outcomes on the other hand, it is necessary to conduct multivariate analyses involving more than these two variables. In our study, we take a multivariate approach to statistical analyses of individual-level data.

Our multivariate approach can be easily demonstrated with the language of direct, indirect, and total effects used in path analysis and structural equation modeling (Alwin and Hauser 1975). In Figure 1.2, we give an unrealistically simple presentation for illustrative purposes. Let X denote the variable of sex (coded as female = 1, male = 0), and Y denote the outcome variable. The bivariate effect of sex, shown as A* in Figure 1.2a, is called the total effect. In Figure 1.2b, this total effect of sex on Y is decomposed into two components: a direct effect (A) and an indirect effect through covariate Z (paths C and B).[7]

[Figure 1.2 about Here]

It should be noted that sex (X) is causally prior to potential covariates (Z) so that Z represents all possible mediating variables between X and Y. In this simple setup, we are interested in the relative importance of the indirect effect of X on Y through Z. When the indirect effect constitutes a large part of the total effect of X on Y, we attain a good understanding of how the total effect of X on Y operates. It is in this context that we say that Z "explains" the effect of X.

Not all potential covariates serve to mediate the total effect of sex (X). To illustrate this point, let A* be negative, signifying that the outcome variable has a lower mean for women than for men. When covariate Z serves to mediate women's disadvantage in Y, we mean one of the following two scenarios:

Both of the following two conditions are satisfied: (1) Z positively affects Y; and (2) Z is negatively associated with sex, with women having a lower mean of Z than men. Or,

Both of the following two conditions are satisfied: (1) Z negatively affects Y; and (2) Z is positively associated with sex, with women having a higher mean of Z than men.

We call (1) the relevance condition and (2) the correlation condition (in either scenario). A covariate may confound the bivariate relationship between sex and an outcome variable only if the covariate is relevant to the outcome variable and correlated with sex.

However, one should not mechanically apply the relevance and correlation conditions. Doing so would lead to the inclusion of too many covariates in a multivariate analysis and erroneous conclusions if the covariates are "pseudocontrols" (Lieberson 1985). It is critical to select covariates that are appropriate for each particular research context. Each chapter that follows addresses a specific question about gender differences in S/E career processes and outcomes and in that context considers a set of covariates that may serve as explanatory factors.

As we discussed earlier, we can easily compute the sex ratio of means (*R*) and sex ratios of odds (*OR*) from regression models with logged mean, and logged odds, respectively, as the dependent variable. In such cases, *R* or *OR* is simply the exponentiated coefficient of the sex variable, $\exp(b_{sex})$. Since sex is a dummy variable coded 0 for male and 1 for female, $\exp(b_{sex})$ is called the female-to-male ratio (or odds ratio) in tables reporting results from multivariate analyses. When women are disadvantaged relative to men for a positive career outcome, b_{sex} is negative, and $\exp(b_{sex})$ is lower than one. As covariates are added, the magnitude of $\exp(b_{sex})$ changes. If the covariates included indeed help explain the total effect of sex on the outcome, $\exp(b_{sex})$ increases in value toward one. The increase in $\exp(b_{sex})$ over its value in the baseline bivariate model (i.e., with no other controls) yields a sensible measure of the extent to which the explanatory variables included in a statistical model "explain" the raw bivariate sex difference in the outcome. However, the size of the change in $\exp(b_{sex})$ depends on what other variables are present in the model. In Appendix B, we describe a method that allows us to estimate the independent contribution of an explanatory variable in accounting for changes in $\exp(b_{sex})$. We use this decomposition methodology in several chapters.

Outline of the Book

This book is divided into two parts. In the first part, we focus on gender differences along the S/E educational trajectory. We begin with academic achievement in science and math in pre-college years in Chapter 2 and then move to the expectation of enrolling in an S/E major in college in Chapter 3. In

Chapter 4, we document gender differences in the likelihood of obtaining an S/E degree at the bachelor's level, paying attention to aggregate transition rates leading to the attainment of S/E degrees as well as individual-level determinants. In Chapters 5 and 6, we are concerned with career transitions following the completion of an S/E degree respectively at the bachelor's level and at the master's level. Beginning with Chapter 7, we shift our attention from career processes to career outcomes, with an empirical portrait of the demographic and labor force outcomes of men and women scientists. In Chapter 8, we focus on the likelihood of geographic mobility among all scientists. Research productivity among academic scientists is the subject of Chapter 9. In Chapter 10, we explore the consequences of immigration on the gender composition of scientists and on the gender differences in labor force outcomes among scientists.

Endnotes for Chapter 1

---

[1] There are no concrete indications, however, that this concern is justified. Of course, part of the reason why the U.S. is able to maintain its large scientific labor force is through immigration. The interaction between gender and immigration is discussed in Chapter 10.

[2] See Keller (1985, 1992) and Harding (1986) for the feminist contention that women bring unique epistemological perspectives to science. See Koertge (1998) for a recent critique and Schiebinger (1999) for a recent review.

[3] For convenience, we sometimes refer to science/engineering as science, and scientist/engineer as scientist, throughout the book.

[4] This statement does not contradict the fact that some excellent samples have been created for the study of women in science. Prime examples are Cole (1979) and Long, Allison, and McGinnis (1993). Common to these studies is a research design that collects data from recipients of doctoral degrees in certain scientific fields.

[5] The excellent study of Citro and Kalton (1989) led NSF to stop using this hybrid definition and to adopt two alternative, education-based and occupation-based, definitions, to be discussed in the next paragraph.

[6] Our discussion ignores sampling variability and thus treats estimated statistics as population parameters.

[7] Since the statistical models may be nonlinear, the indirect effect cannot always be computed as the simple product of the two direct effects C and B.