INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

UMI

A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor MI 48106-1346 USA
313/761-4700 800/521-0600

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

A MICROECONOMIC EXAMINATION OF THE RELATIONSHIP BETWEEN OCCUPATIONAL GENDER SEGREGATION AND TASTE DIFFERENCES

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

DOCTOR OF PHILOSOPHY

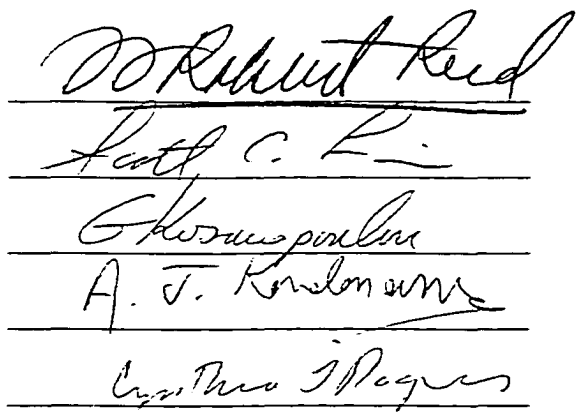
By

James R. Wilbanks
Norman, Oklahoma
1998

A MICROECONOMIC EXAMINATION OF THE RELATIONSHIP BETWEEN OCCUPATIONAL GENDER SEGREGATION AND TASTE DIFFERENCES

A DISSERTATION
APPROVED FOR THE DEPARTMENT OF ECONOMICS

BY



Acknowledgements

The number of persons to be acknowledged are numerous. I would first like to thank the members of my committee. I would especially like to thank my chairman, W. Robert Reed, without whom this work would not have been completed. I am extremely appreciative of the consistent and fair treatment. At times, he was harsh and his criticism biting, but in retrospect, it was well deserved.

Another person that deserves my thanks and gratitude is David Merrell. He has always been available for discussion throughout this work. In addition, he willingly sacrificed his precious time to read earlier drafts, and provide insightful comments and criticism.

Finally, and most importantly, I would like to thank my wife Laura and my son Zachary, both of whom provided motivation for me to continue my work. Zachary is too young to know the impact he has had upon me, but suffice it to say that without him, I might have never completed this project. And for Laura, who always seemed to know when to use the stick and when to use the carrot, I owe eternal thanks. Unknown to most, she has been the driving force behind the completion of this work and without her, I would be a ship adrift on the seas.

To the memory of my Grandmother and Grandfather, whose pride in me was surpassed only by their support and love.

TABLE OF CONTENTS

List of Tables

List of Figures

Abstract

This work examines the impact of taste difference between genders for non-pecuniary work characteristics on occupational outcome. This is accomplished by estimating reservation wages for workers across several job categories. Additionally, wage distributions are estimated in the occupations. This allows for the generation of a distribution of workers across jobs. By matching workers with jobs based on only reservation wages and the distribution of wages, the resulting distribution of workers is unaffected by any employer controlled factors. Several sets of results are generated based on different assumptions concerning the amount of information used in matching workers with jobs. Each of these sets of results is then examined for gender segregation and also compared to the segregation observed in the labor market. In general, the results here find some evidence that gender based taste differences matter in the job matching process, but this evidence is modest. Further, there is no evidence to support the hypothesis that these taste difference explain the segregation patterns in the workplace.

Chapter 1

INTRODUCTION

Discrimination is defined as singling out for unfavorable treatment (Webster's, 1992). Often this singling out is based on some characteristic of individuals. This characteristic can be anything that is measurable, either objectively or subjectively. Examples of measurable characteristics include effort level, race, intelligence, gender, physical ability, religion, prior performance, and physical appearance.

The primary focus of public policy in this area has been on discrimination that is based on characteristics that individuals cannot change, such as race or gender. The reason for this focus is that it is seen as unfair to discriminate against a person based on something over which they have no control. Conversely, it is viewed as acceptable to discriminate based on characteristics that individuals can control. For example, it is rather unlikely that a student could successfully claim they were being treated unfairly in the assignment of course grades if the characteristic on which grades were assigned was performance in the course, even though this is discrimination.

Discrimination can take many different forms. One form of discrimination is segregation. To segregate is to separate from others of a group, to isolate (Webster's, 1992). Segregation is often a manifestation of discrimination in which the unfavorable treatment is separation of the group or individual being targeted from other groups or individuals.

However, the existence of segregation does not, by itself, indicate the practice of discrimination. This is because segregation can occur for reasons other than just discrimination. For example, consider the seating of customers in a college dormitory cafeteria. Customers typically will seat themselves in a segregated manner. That is, males tend to sit with males, women tend to sit with other women, black customers tend to sit with black customers, and white customers tend to sit with other white customers. When this happens, customers are clearly being segregated. However, the source of the segregation is not the cafeteria operators, it is the customers themselves. This example demonstrates how segregation can occur without the presence of discrimination.

This study will address segregation in labor markets. More specifically, it will focus on occupational gender segregation. Occupational gender segregation occurs when job categories are dominated by one gender. For example, in 1989 over 90% of those working in the occupation nursing were females. Conversely, in the same year less than 10% of those employed as engineers, were females (U.S. Department of Labor, 1990). These two job categories were highly gender segregated in 1989.

However, as discussed above, the existence of segregation is not a sufficient condition to conclude that discrimination is present. This is due to the fact that segregation can be caused by many different factors. Discrimination is only one of the potential causes of occupational segregation. Other potential factors include differences in human capital, differences in tastes and preferences, differences in alternatives to work, and differences in costs associated with work.

This study will focus on the portion of occupational gender segregation that is attributable to differences in tastes and preferences between males and females. Specifically, the emphasis of this work will be to estimate the amount of occupational gender segregation that is caused by differences in tastes for non-wage characteristics between genders. The primary question this work attempts to address is 'Do gender differences in tastes for work explain any of the observed gender segregation, and if so, how much?'. To accomplish this, a distribution of workers across occupations must be generated.

The next two chapters will focus on the development of a labor market model that will allow for estimation of the required distribution of workers. Chapter Four will discuss the estimation methods used to arrive at this distribution. Chapter Five will focus on the data employed in the estimation process while chapter Six will present the results of the estimation. Finally, chapter Seven will be a discussion of the conclusions of the study and extensions for future work. The remainder of this chapter will be devoted to a discussion of relevant literature.

Virtually any research on the topic of discrimination stems from the work of Gary Becker (1971). In his work, Becker modeled the labor market such that discrimination was manifest as wage differentials. These wage differentials are the result of 'tastes' for discrimination against some minority. These preferences could arise from any of three different agents, the employer, other employees, or customers.

Any of these actors that have preferences not to associate with the minority would realize a negative change in their utility if they did associate with the minority.

Accordingly, this impact on any agent's utility would have to be considered in their decisions concerning employment, work, or conducting business. The agent with the distaste for associating with the minority would require compensation to induce them to make such an association (Becker, 1971).

The discriminatory employer would perceive lower productivity for the minority employee due to their discrimination taste factor. This, in turn, results in a lower wage payment for the minority. A discriminating co-worker would expect a higher wage payment to account for the disutility of working with a minority. This results in a lower net productivity for the minority worker. A customer with discriminatory preferences would expect to pay a lower price for the firm's product as compensation for the loss of utility from associating with the minority. This lowers the value of the minority worker's output, thereby decreasing his or her productivity (Becker, 1971).

The result of the discrimination, regardless of the source, is that a minority worker has lower productivity, measured net of discrimination, than their majority counterpart. Thus, according to Becker (1971), wage differentials in the labor market are the result of discrimination on the part of one of three agents. Becker refers to this wage differential as the discrimination coefficient.

Having quantified discrimination in monetary terms, Becker moved further in his work to address discrimination in the long run. He showed that, in the long run, firms that practiced discriminatory hiring and compensation would be driven out of competitive markets (Becker, 1971). Firms hiring minority workforces could have

lower wage payments since the minority workers would face a lower labor demand. The smaller wage bill would then allow the firms hiring a minority workforce to charge lower prices for their products. In competitive markets, firms with higher prices are driven out. Thus, in the long run labor market discrimination would all together disappear due to the forces of competition (Becker, 1971).

However, Becker's theory could explain long term discrimination in monopolistic markets. With no competitors to drive a firm out of the market, a monopolist is free to discriminate without fear of market consequences. It is unlikely, however, that the perceived level of discrimination in today's labor market can be attributed solely to discrimination in monopolistic markets.

In a refinement and addition to Becker's work, Kenneth Arrow presented possible explanations of long run discrimination (Arrow, 1973). Arrow showed that at least three different causes exist that can explain long run discrimination. The first is that indifference surfaces and opportunities are nonconvex. These nonconvexities will lead agents to create niches for themselves. These niches then propagate discrimination.

The second factor that leads to long run discrimination is the cost of adjustment (Arrow, 1973). It is argued that an employer would not be willing to switch from an all majority to an all minority workforce to save a small amount in wage payments. This is due to the fact that there is a sunk cost of training employees and this would be lost by firing all majority workers. Employing some minority workers would increase the wage payment and thus is not in the best interest of a

profit-maximizing firm. Thus, these costs of adjustment could explain discrimination in the long run.

The third reason that Arrow provides to explain long run discrimination is that of imperfect information (Arrow, 1973). That is, if employers have imperfect information concerning prospective employees, which they always do, then employers might use some characteristics of the applicant to infer some information. These characteristics could include race, creed, color, or gender. This information would be based on what the employer believed the 'average' characteristics were for that particular minority. Thus, we can see that there are possibilities that discrimination can exist in the long run even in light of the opposing competitive forces outlined by Becker's model.

Another model that may help to explain the existence of discrimination in the long run is that of Phelps (Phelps, 1972). The model of Phelps' shows the existence of 'statistical' discrimination. One example of this type of discrimination is when an employer infers information from a non-changeable attribute of a prospective employee. This is one of the explanations offered by Arrow for the existence of discrimination in the long run (Arrow, 1973).

Another example of statistical discrimination described by Phelps can occur when a skills test (intended to predict future productivity) is administered to prospective employees. If this skills test estimates a minority person's future productivity less predictably the employer could be less likely to hire the minority

worker (Phelps, 1972). Aigner & Cain improve upon this argument by including the employer's risk factor (Aigner & Cain, 1977).

All of these models of discrimination have one thing in common. They all explain discrimination via the demand for labor. The discrimination is practiced by the employer or potential employer, regardless of the source of the discrimination. There is good reason for this.

There is no discrimination against workers via the supply of labor. This is because individuals do not single themselves out for unfavorable treatment; this singling out must be done by some other agent. Thus, no discrimination can be explained via labor supply factors.

However, what many people perceive as discrimination may actually be caused by labor supply decisions. As indicated above, individuals may make choices that result in outcomes that give the appearance of discrimination. For example, suppose group A individuals tend to prefer working in a certain job category more than group B individuals. Is discrimination occurring if more employees in that occupation are from group A than B?

If the two groups are of comparable size, it is unlikely that discrimination is occurring. Rather, the labor market is quite likely operating in an efficient manner. This is due to the fact that the individuals are being distributed across occupations in a manner consistent with maximizing their utility according to their preferences.

Because of the impracticality of discrimination in labor markets resulting from the supply side, supply side factors should instead be thought of as causing

differential outcomes rather than discrimination. It is to these supply side factors that the discussion now turns. The first serious attempt to explain differential outcomes with supply side causes was made by Solomon Polachek (Polachek, 1975).

Polachek developed the human capital approach to explaining differential outcomes. That is, individuals acquire human capital (education or training) in order to maximize lifetime utility. The reason this leads to differential outcomes is that different minorities might acquire differing amounts of human capital. The differing amounts of human capital, by group, would then be expected, in turn, to result in differences in the distribution of workers across occupations and wages (Polachek, 1975).

Polachek applies this model to explain why women would choose a smaller investment in human capital than men would if they expect to experience intermittent participation in the labor market (Polachek, 1981). That is, since many women expect to leave the labor force to bear and rear children they should invest in less education and training than comparable males. This is caused by the reduced time over which benefits from the investment can be gained.

As indicated previously, the focus of this research is occupational gender segregation. Segregation is a result of many factors including discrimination. The specific factor that is studied here is segregation caused by differences in tastes and preferences. This type of segregation may be referred to as self-selection since it occurs due to individuals maximizing their utility. By doing so, they choose to enter

a certain occupation rather than some others. This discussion now turns its focus specifically to gender segregation in the labor market.

There is a considerable amount of existing research on the phenomenon of occupational gender segregation. Most of this work can be grouped into three broad categories. The first, and largest, examines the different causes of segregation as well as the amount of segregation present in the labor market (Blau & Ferber, 1992; Reskin & Hartmann, 1986).

The second group of work examines policies aimed at correcting the perceived problem. Some of the policies that are discussed include comparable worth (Filer, 1989) and affirmative action legislation and enforcement (Leonard, 1989). Comparable worth is the idea that two workers in distinctly different jobs should be paid equally if each job contributes a similar value to a firm's output. Affirmative action refers to government policy of preferential treatment of certain minorities in hiring and promotion decisions in an effort to achieve non-discriminatory outcomes. Some works contribute to both of the first two categories of research (Reskin & Hartmann, 1986; Michael et. al., 1989).

The third category is research that examines a specific cause of segregation and attempts to estimate the amount of segregation that occurs because of that factor. Some of the specific causes that have been examined include taste differences (Daymont and Andrisani, 1984; Filer, 1986; and Gupta, 1993), and differences in human capital investment (Polachek, 1981; England, 1982). The works that explore taste differences find evidence that the amount of segregation attributable to this

cause may be significant. Results from the human capital studies are not as clear as those from the works on taste differences.

It is to this third category that this paper will contribute. Some of the existing literature on gender segregation has found differences in taste as a minor factor in the explanation of total segregation (Blau and Hendricks, 1979; Blau and Ferber, 1992). As stated above, evidence has also been presented to support taste differences as a significant cause of segregation. A brief discussion of the previous works that have examined the relationship between segregation and gender based taste differences will show this and bring into focus the contribution of the current research.

The work of Daymont and Andrisani explores the role of tastes as they affect the choice of occupation and, in turn, how the occupational choice affects the gender pay gap (Daymont and Andrisani, 1984). The authors employ data from the National Longitudinal Survey for the High School Class of 1972 to estimate earnings equations for males and females separately. A decomposition analysis is then undertaken in an effort to determine what amount of wage differences are caused by various factors including preferences for various types of jobs. The authors show that between six and 27 percent of the male-female wage gap can be explained by differences in occupational choice that are attributable to taste differences (Daymont and Andrisani, 1984).

A study by Filer used a unique data set from the personnel files of a management-consulting firm covering roughly 3,800 individuals (Filer, 1986). One key characteristic of the data employed was that it contained information about the

individual's traits, which are expected to represent their tastes. Logit model estimates are formed of the probability of individuals choosing different job types relative to choosing a managerial job.

These estimates are first formed using demographic information but no measures of preferences. The logit model is then re-estimated adding in the preference variables. The two sets of estimates are then compared as to their predictive ability (Filer, 1986).

The first model, using only demographic variables, is able to correctly predict the occupation of an individual 40% of the time. Adding the preference variables into the model increases the predictive ability of the model. The second set of estimates can correctly predict occupational outcome in 46% of the cases. While this difference is not enormous, it is sizable enough to warrant a conclusion that tastes do matter in choice of occupation and therefore cause at least some occupational gender segregation (Filer, 1986).

In another work on the topic of the role of tastes on occupational gender segregation, Gupta uses National Longitudinal Survey, Youth Cohort (NLSY) data to estimate the probability of workers choosing different broad job categories (Gupta, 1993). An estimate is also formed of the influence of employer selection on the probability of different individuals working in these categories. These latter estimates are not discussed here since the focus of this work is segregation caused by taste differences. The results of the estimates of workers' job category choice indicate that females tend to have a higher probability of choosing to work in female dominated

occupations than males while males have a higher probability of choosing to work in male dominated occupations (Gupta, 1993). These results are similar to those of Filer about differences, based on gender, in worker's probability of choosing different occupations.

The primary shortcoming of these works is that they do not attempt to directly estimate the amount of segregation resulting from taste differences. Differences in taste for job characteristics have been examined either in terms of explaining the gender pay gap or in terms of determining the probability of job choice. The present work will attempt to estimate the amount of segregation actually caused by taste differences between genders. Estimating the amount of segregation due to differences in preferences may have significant implications for policy directed at occupational segregation and its elimination. An examination of the causes of segregation shows why this work can have such policy implications.

The causes of segregation can be placed into two different groups: voluntary and involuntary. The causes of observed segregation that are involuntary include, but are not limited to, discrimination in hiring procedures, job typing (Reskin and Hartmann, 1986; Blau and Ferber, 1992), and statistical discrimination[1] (Phelps, 1972; Aigner & Cain, 1977). Causes of segregation that are voluntary include differences in human capital investment (Polachek, 1981; England, 1982) and self-

[1] There is some debate over whether statistical discrimination is actually discrimination at all. If individuals are compensated based on the average productivity for a group to which they belong, the employer is not discriminating against all individuals in that group. Those in the group with productivity above the group average are being discriminated against while those with productivity

selection or differences in tastes (Daymont & Andrisani, 1984; Filer, 1986; Gupta, 1993). Only segregation that is involuntary should be of interest from a policy perspective given the classic assumption of self motivated behavior. That is, individuals will self-select into, or away from, an occupation only if doing so is welfare maximizing.

For this analysis, the first issue is how to quantify taste differences across individuals. The method employed in this study to analyze preferences for the job characteristics of different occupations is by comparison of reservation wages. A reservation wage will provide the necessary information concerning tastes since it is determined independently by individuals making assessments about a job and its characteristics based solely upon subjective preferences.

Specifically, the lower a person's reservation wage for a certain job, the more they would prefer that job. Examining an individual's set of reservation wages for several different jobs reveals in which of those jobs that individual would most prefer to be employed. Clearly then, the first task to be accomplished is to obtain reservation wages in several job categories for many individuals. This study obtains sets of reservation wages using the 1979 National Longitudinal Survey, Youth Cohort survey data.

The next major issue in this study is generating a distribution of workers based on their reservation wages. This is accomplished by forming estimates of the

---

below the group mean are the beneficiary of higher wages. On average, however, the group members are compensated according to their productivity.

likelihood of a worker accepting an employment offer in each job category. Workers are then matched with jobs resulting in a distribution of workers across jobs. Finally, the gender composition of each job category will be measured. This distribution will then be compared to the actual distribution of workers across job categories, allowing for an inference to be drawn about the importance of individual preferences in occupational segregation.

Several different methods for generating the needed distribution of workers across occupations are developed in this work. These methods differ in the assumptions underlying the approach and the amount of available information that is used in forming the distribution. A total of five different distributions of workers are generated.

The results of these different distributions are mixed. Two of the methods used result in a finding that a significant amount of segregation is caused by self-selection. The other three methods generate the opposite result that no segregation is caused by self-selection. These differing outcomes allow for no definitive conclusion concerning the amount of observed occupational gender segregation is caused by workers self-selecting into occupations based on their tastes and preferences.

The primary contribution of this work then is the development of a method for generating a distribution of workers across job categories. This contribution, along with future refinements and enhancements, should better equip future researchers to address the phenomenon of gender segregation in the workplace. Specifically, these

techniques should enable future work to better estimate the amount of occupational gender segregation that is caused by self-selection.

Chapter 2

SEARCH

The task undertaken in this and the next chapter is the development of a labor market model that shows a direct connection between workers' preferences and occupational outcomes. This chapter will develop the model by which workers search for employment. In so doing, it will also demonstrate the existence of a relationship between worker's tastes and their search activity. Specifically, a connection will be shown to exist between a worker's subjective evaluation of the non-wage characteristics of jobs and that worker's search activity.

Chapter 3 will then show the relationship between workers' search and occupational outcomes. A brief discussion of the background of search theory will begin this chapter. Next a formal model of worker search will be developed. The chapter will conclude with a discussion of empirical issues of interest to the search model.

The first formal treatment of search theory was that of Stigler (1961). In this work, Stigler examines the role of search as a response to price dispersion and consumer ignorance concerning the location of the best price. Stigler considers the consumer's problem of determining the appropriate number of searches. This determination is based on the expected marginal benefit and expected marginal cost from additional search (Stigler, 1961).

Consumers search for, and accumulate, many different prices for a particular good. After searching for an optimal number of prices the consumer purchases the

good at the lowest price encountered (Stigler, 1961). This is frequently referred to as the fixed sample size search method since the number of searches is fixed once it is determined.

Soon after his original work, Stigler applied this search model to the labor market (1962). As before, Stigler considered the problem of determining the optimum number of searches based on marginal benefit and marginal cost of each additional search. Stigler then extended the analysis to explore the relationship of several factors thought to affect the benefits and costs of search. Accordingly, these factors also were expected to affect the amount of optimal search (Stigler, 1962).

For example, Stigler explained that the benefit from one additional search increases with the length of expected employment. This, in turn, results in higher wage offers being accepted which implies fewer low wage offers are accepted. The end result is an observed wage distribution with a smaller variance (Stigler, 1962). Anecdotal evidence was presented to support this claim using Census data from 1940 and 1950.

However, the fixed sample size aspect of the Stigler labor search model may not be consistent with the actual functioning of the labor market. Job applicants are rarely able to accumulate some optimal number of job offers, then decide among them. Typically, when a potential employee is offered a job, that employee has some finite, and usually relatively short, amount of time to make a decision about the offer. In many cases, the employee may have as little as a few moments.

An alternative model of the search process, that may better explain the market for labor, is based on an optimal stopping rule. McCall (1970) formalized this model of search in the labor market. The basis of this model is that a worker conducts searches in a chronological order.

When an individual receives an employment/wage offer, they determine whether or not they are willing to accept the employment at the offered wage. If they are not willing to accept the wage, the search is continued. If the worker is willing to accept the offered wage, however, they accept the employment offer and discontinue further search (McCall, 1970). This is frequently referred to as the sequential search model. Mortensen (1986) shows that sequential type search dominates non-sequential search since the expected present value of future income is higher with sequential type search.

To demonstrate this, consider the following example. According to the Stigler model, a searcher has a pre-determined number of searches that is optimal for them. Suppose that upon the first search, the employee finds a wage in the upper tail of the known distribution of wages. According to Stigler's model, the employee should continue to search. The sequential search model improves search theory by recognizing the fact that when the employee finds a wage with which he or she is satisfied, they stop searching and accept employment.

This is an optimal response on the part of the searcher. If search continues after a high wage has been found, costs continue to increase. At the same time, expected benefits are unlikely to increase since a relatively high wage has already

been found. Thus, the searcher would be lowering his or her expected present value of future income by continuing to search.

In an effort to address concerns with the sequential search model, Gastwirth examines the typical assumption of full information concerning the distribution of prices (Gastwirth, 1976). He shows that if a searcher uses incorrect information about the distribution of prices, the resulting sequential search may not be optimal. In an effort to further address this issue Rothschild examines the properties of optimal search if the distribution of prices is unknown to the searcher (Rothschild, 1974). His results indicate that optimal search method is qualitatively similar whether the individual is searching over known or unknown distributions.

In spite of this, there remain specific examples where fixed sample size type search dominates sequential search. In their 1982 paper, Manning and Morgan demonstrate that fixed sample size search is optimal if the specific strategy is to determine sample size for each period and the costs of search after the first period are prohibitively high (Manning and Morgan, 1982). An example is the job search by a new Ph.D. who is searching for employment in academia.

In a later work, Morgan and Manning show that using a mix of sequential search and fixed sample size search has a higher expected payoff than either pure search strategy (Morgan and Manning, 1985). However, this strategy collapses to sequential search if the time period considered is assumed to be short enough such that only one wage can be found in any time period. As will be discussed in chapters four and five, the jobs over which workers search in this model are of the low-skill,

low-pay variety. In such job categories, it would seem a reasonable assumption that workers have little time to accept or reject an offer, possibly a matter of minutes.

This assumption of time periods sufficiently short that only one offer can be generated results in workers searching in a purely sequential manner. Accordingly, the search model developed in this chapter will draw most heavily upon the sequential search models of McCall (1970) and Mortensen (1986). However, the model used herein will be a somewhat simplified version.

Clearly, workers search for employment/wage offers due to the existence of wage differences. These differences may exist, even for identical jobs. Prior to being offered employment, workers cannot determine the wage that a particular firm will offer. However, workers are assumed to know the distribution of wages.

As stated above, the search that workers undertake is assumed to be purely sequential. That is, when a worker is offered a job, that worker either accepts or rejects the offer immediately with no possibility of recall. These assumptions merely serve to simplify the search model.

The search that workers undertake is costly. There are two different costs typically associated with an individual searching for employment. The first is the direct cost of time and resources spent in the search process. The second cost is the indirect cost of forgone earnings during search. These costs are assumed to be non-zero, and quite substantial, so workers do not search infinitely.

The costs of search, and more importantly, the sequential nature of the search process, necessitate that each worker form a decision rule for accepting or rejecting

any potential job offer. This decision rule will incorporate the offered wage, expectations about the job's characteristics, and expectations about the characteristics and wages of other potential job offers. This decision rule is the first matter to be addressed by this model.

For a worker to accept any job offer, the offered wage must meet or exceed some threshold level. This threshold is referred to as a reservation wage. The reservation wage is the lowest wage at which a person will accept employment in the given occupation. Necessarily, the reservation wage is also the wage at which the individual is indifferent between accepting or rejecting the job offer.

If a worker is offered a certain job at some specified wage, that worker will accept the offer if the wage is at least as high as the reservation wage. If the wage is lower than the reservation wage, the worker will reject the offer. Thus, a worker's decision to accept or reject an offer is based only on his or her reservation wage for that job and the offered wage. The next matter to be considered by this model is an individual's reservation wage for different jobs and what determines these reservation wages.

A formal model of the worker and their search will aid in the development of reservation wages and their determinants. The models developed here are similar to the models of Mortensen (1986), Devine and Kiefer (1991), and Khandker (1988). Before presenting the model, listing some assumptions used for simplification may be helpful.

First, individuals maximize utility by maximizing the expected present value of their future stream of income discounted at interest rate ρ. Workers are assumed to search over J different job categories where $J \in \aleph$. Each job offer is then characterized by not only the wage but also the job itself. To account for this, jobs within a category will be assumed identical in every way except wage. Different jobs, however, can vary in non-wage characteristics as well as the wage. Thus, a job offer can be characterized by the wage and the non-wage characteristics of the job[2].

With the worker searching over many different jobs their instantaneous utility can be written as

$$U_i = \begin{cases} b - c & \text{, if unemployed and searching} \\ w_j + v_j & \text{, if employed in job j for } j = 1,2,\ldots,J \end{cases} \qquad (1)$$

where b represents the income and utility received while unemployed and c is the cost of search. So b - c is the net cost of search and is unrestricted in sign. Also, w represents the wage which is a random variable determined by $w_j = \mu_j + \varepsilon_j$ where μ_j represents the mean of the wage distribution and ε_j is a white noise error term. The term v_j represents the dollar value of the utility derived from the non-wage characteristics of job j. It is clear that the worker's utility is assumed to be additive in wage and non-wage attributes of a job.

Individuals search in order to generate job offers. The arrival of offers is a random event assumed to be characterized by a Poisson process with parameter δ.

---

[2] Non-wage characteristics of a job include, but are not limited to, work environment, hours, part/full time, type of work (manual labor, clerical, etc.), and distance to work.

Also, each job offer is assumed to be an independent realization from a known and well behaved wage distribution. Finally, each job is assumed to be infinitely lived, meaning no separations.

Having listed the above assumptions, the model can now be presented in a relatively simple manner. When a worker receives a job offer that worker then has the following value function for accepting an offer in job category j,

$$V_j^E = \frac{w_j + v_j}{\rho}. \tag{2}$$

This shows that the value to the worker of accepting the job offer is the expected present discounted value of the infinite stream of income from the job.

Correspondingly, each individual also has the value function for remaining unemployed in the current period as follows:

$$V^U = \frac{(b-c)}{\delta+\rho} + \frac{\delta}{\delta+\rho}\sum_{j=1}^{J} P_j E\left\{\max\left[V_j^E, V^U\right]\right\} \tag{3}$$

where P_j is the probability that a given job offer is in job category j and the sum of all P_j is equal to 1. The first term in equation 3 represents the net cost of search during the current period. The second term in equation 3 represents the probability of finding a job offer multiplied by the expected discounted value of following a policy that maximizes the worker's utility.

The J employment equations plus the one unemployment equation can be solved simultaneously for reservation wages. As previously stated, the worker's reservation wage is the wage at which that worker is indifferent between accepting

employment and remaining unemployed. That is, the reservation wage, w_j^r, solves the equation $V^U = V_j^E$. As equation 2 indicates this can be restated as the following:

$$V^U = V_j^E = \frac{w_j^r + v_j}{\rho}. \tag{4}$$

Some minor rearranging of equation 4 reveals

$$w_j = \rho V^U - v_j = w_j^r. \tag{5}$$

$E\{\max[V_j^E, V^U]\}$ from equation 3 above can be rewritten as $V^U + \int_{w_j^r}^{\infty}(V_j^E - V^U)dF_j(w)$. Equation 3 then becomes

$$V^U = \frac{(b-c)}{\delta+\rho} + \frac{\delta}{\delta+\rho}\sum_{j=1}^{J} P_j\left\{V^U + \int_{w_j^r}^{\infty}(V_j^E - V^U)dF_j(w)\right\}. \tag{6}$$

Some simplification of equation 6 results in the following,

$$\rho V^U = (b-c) + \delta\sum_{j=1}^{J} P_j\left\{\int_{w_j^r}^{\infty}(V_j^E - V^U)dF_j(w)\right\}. \tag{7}$$

Substituting equation 2 into equation 7 yields

$$\rho V^U = (b-c) + \delta\sum_{j=1}^{J} P_j\left\{\int_{w_j^r}^{\infty}\left(\frac{w_j + v_j}{\rho} - \frac{\rho V^U}{\rho}\right)dF_j(w)\right\}. \tag{8}$$

which can be simplified to

$$\rho V^{U} = (b-c) + \frac{\delta}{\rho}\sum_{j=1}^{J} P_j \left\{ \int_{w_j^r}^{\infty} \left(w_j + v_j - \rho V^{U} \right) dF_j(w) \right\}. \tag{9}$$

Substituting equation 5 into equation 9 and simplifying yields

$$\rho V^{U} = (b-c) + \frac{\delta}{\rho}\sum_{j=1}^{J} P_j \left\{ \int_{w_j^r}^{\infty} \left(w_j - w_j^r \right) dF_j(w) \right\}. \tag{10}$$

Substituting equation 10 into equation 5 results in the following,

$$w_j^r = (b-c) + \frac{\delta}{\rho}\sum_{j=1}^{J} P_j \left\{ \int_{w_j^r}^{\infty} \left(w_j - w_j^r \right) dF_j(w) \right\} - v_j. \tag{11}$$

This result, which is also shown by Khandker (1988), indicates that the reservation wage in job j is only affected by factors exogenous to job j and the utility offered by the non-wage characteristics of job j. The utility offered by these non-wage characteristics of a job negatively affects the reservation wage. That is, the less desirable a worker finds a particular job, the higher the individual's reservation wage for that job will be. The exogenous factors that affect w^r through the term ρV^u are the interest rate (ρ), the net cost of search (b - c), the vector of probabilities of different job categories (P), the vector of mean wages (μ), and the vector of utilities from non-wage characteristics of jobs j (V).

In this model, workers are assumed to search over many jobs. Additionally, the search that workers undertake is not specific to job categories. That is, workers are assumed to search over all job categories simultaneously. Accordingly, they face a distribution of wages that combines all of the wage distributions of the jobs in

which they search. The result presented here, that the reservation wage in a particular job is determined exclusively by exogenous factors and the non-wage utility of that job, follows directly from the assumption of search over many job categories and more importantly, the lack of information concerning job specific wage distributions.

If workers had information concerning the distribution of wages in specific job categories, this information would certainly affect their reservation wages. However, the influence would not necessarily be consistent across occupations. The resulting estimates of a given individual's reservation wages would differ for reasons other than the worker's tastes for the non-wage attributes of the jobs. The result then is that the direct link between tastes and reservation wages is not present and workers cannot be matched with occupation based on preferences. Thus, the assumption that workers have no information concerning wage distributions in specific occupations is critical.

The reservation wage that is developed with the above model, shows how these reservation wages differ across jobs for the same individual. By comparing any two reservation wages, it is clear that the term ρV^U is constant across all job categories for one worker. Thus, the only difference between the worker's reservation wages is the utility associated with the non-wage characteristics of those jobs. To see this more clearly, consider the measure d_{ijk}, which shows the difference in the i^{th} individual's reservation wages for jobs j and k (McCue and Reed, 1996). This is found as the following:

$$d_{ijk} = r_{ij} - r_{ik} = v_{ij} - v_{ik}. \qquad (12)$$

This measure, if positive, quantifies, in dollar terms, the amount by which the non-wage characteristics of job j are preferred by individual i to the non-wage characteristics of job k. If d_{ijk} takes a negative value, then it shows the dollar amount by which job k is preferred to job j in non-wage characteristics. The only difference in reservation wages for an individual is the difference in the utility associated with the non-wage characteristics of each job since workers search over many different jobs.

It is possible that outside factors may affect the individual's preferences for the attributes of the job. For example, suppose an individual has previously observed discrimination in a certain occupation. This person could then include the discrimination as an attribute in their assessment of that job category. Clearly, differences in reservation wages would not then be attributable solely to differences in the preferences for jobs.

In the model above, as well as empirically, there is no way to capture and remove the effect of external factors on an individual's tastes for a job. Consequently, the only remaining solution is to assume that external factors impact a given individual's preferences for the attributes of different jobs in a similar manner. That is, the effect of external factors on reservation wages are assumed constant across jobs. Given this assumption, any differences in reservation wages can be attributed solely to differences in preferences.

Since the objective of this work is to examine the relationship between gender segregation and taste differences, a connection must be drawn between preferences

and occupational outcomes. The connection that the above model provides is between tastes and reservation wage. The more distasteful an individual finds a particular job, the higher will be their reservation wage for that job. The remaining connection of how reservation wages affect occupational outcomes and thus gender segregation will be presented in the discussion of matching workers with a specific job in the next chapter.

With the relationship between tastes and reservation wage now clear, a discussion of the nature of reservation wages follows. For the current work, the most important determinant of a worker's reservation wage is the worker's perception of the characteristics of a job. Non-wage characteristics that affect the worker's utility of working in a particular job include, but are not limited to, work environment, job safety, type of work, job status, and perceived 'gender type' of the job.

However, there is no way to quantify the specific amount of utility a worker attaches to the non-wage attributes of any job. What is needed is a way to estimate reservation wages. Obtaining estimates of an individual's reservation wage for different jobs will, in turn, provide a relative measure of that person's level of desire to work in one occupation versus another occupation.

The method used to estimate reservation wages is presented in Chapter 4. The current model is only used to develop reservation wages. Accordingly, a discussion of the exogenous factors that affect reservation wages can be presented within this context. These exogenous factors can then be used to estimate reservation wages. Some simple comparative statics provide a useful starting point for these

relationships. With some rearranging, the following partial derivatives can be found from equation 11 above:

$$\frac{\partial w_j^r}{\partial b} = \frac{\rho}{\rho + \delta P_j[1 - F(w)]} \quad \in (0,1) \tag{13}$$

$$\frac{\partial w_j^r}{\partial c} = \frac{-\rho}{\rho + \delta P_j[1 - F(w)]} \quad \in (-1,0) \tag{14}$$

$$\frac{\partial w_j^r}{\partial \mu} = \frac{\delta[1 - F(w)]}{\rho + \delta[1 - F(w)]} \quad \in (0,1) \quad \forall j \tag{15}$$

Equation 13 shows that an increase in non-work income or the value of leisure would increase a worker's reservation wage[3]. Equation 14 shows that a decrease in the cost of search would also increase the reservation wage. Stated differently, an increase in the utility of not working, or a decrease in the cost of search, would cause a fractional increase in the worker's reservation wages. This result is logical since the reservation wage is the wage at which the value of accepting employment in a job equals the value of remaining unemployed in the current period.

An increase in the utility gained from being unemployed increases the value of remaining unemployed. Accordingly, the reservation wage, for all jobs, increases so that the value of accepting employment equals the value of remaining unemployed. Similarly, a decrease in the cost of search would increase the value of remaining unemployed since it is less costly to remain unemployed and search for an additional period. This in turn results in higher reservation wages in all J job categories.

Also, as equation 15 indicates, an increase in the mean wage would cause a fractional increase in that worker's reservation wage for all jobs. This result is clearly seen by considering what could cause an increase in μ. The clearest example of a factor that would increase μ is an increase in a worker's level of human capital. If a worker increases their human capital, their marginal productivity then increases. In exchange for this increased productivity, the worker will expect a higher wage. Thus, the worker raises their reservation wage.

This analysis provides some insight as to what variables would be useful in estimating reservation wages. Any variable that affects the net costs of search or the mean of the wage distribution, faced by a particular worker, should provide information about that individual's reservation wages. Therefore, a brief discussion of variables that affect these factors will be presented.

Consider first what variables might affect the net cost of search. In the above model, workers are assumed to be unemployed while searching. In reality, however, workers may be concurrently employed and searching for different employment. As such, whether a person is employed will affect their net cost of search by raising the term b in the above model. This results in a corresponding increase in reservation wages.

An individual with a high non-wage income is more likely to be out of the labor force (neither working nor searching for work). Accordingly, such an

---

[3] Pencavel (1986) notes that the value of home production affects the reservation wage. This is clearly captured by a worker's non-work income.

individual is likely to have a higher reservation wage because of the higher non-wage income. Also, if a person is currently enrolled in school, they will likely have a high non-work income since they can afford to commit a non-trivial amount of time to an activity that does not generate current income. This non-work income can take many forms including parental support, spousal support, or the borrowing of money (i.e. student loans). Both of these situations imply a higher value of non-work income and result in higher reservation wages for that individual.

Another set of factors that affect the net cost of search is government assistance (welfare). If an individual receives welfare payments, the effect of these payments is clearly an increase in that individual's non-working income. Additionally, a person that is likely to accept welfare payments if unemployed would have a higher non-wage income when unemployed and searching for work. The increase, or potential increase, in non-work income from accepting welfare results in higher non-wage income. This in turn, results in a higher reservation wage.

The first factor affecting the mean of an individual's wage distribution has already been indicated. That is an individual's level of human capital. Variables that can be thought of as indicating the level of human capital of an individual include: age, highest grade completed in school, whether the person has graduated from high school/college, the type of program/degree in high school/college, previous work experience, and an individual's general knowledge, or intelligence, as assessed by an examination. An increase in any of these indicators of human capital should lead to an increase in the reservation wage.

There are other factors besides human capital that affect the mean wage an individual faces. Possibly the most important of these factors is the environment in which an individual is searching for employment. Indicators of the environment of search include the level of unemployment in the area, the local level of poverty, and local income levels.

An increase in the unemployment level indicates an excess supply of labor. As such, increasing unemployment will, ceteris paribus, be associated with a decrease in the mean of the wage distribution facing a worker. As equation 21 shows, this will cause a decrease in that worker's reservation wage. The result is that higher levels of unemployment imply lower reservation wages.

The amount of poverty on a local level will have a similar effect on reservation wages. This result becomes clear by considering the implications of poverty. A person living in poverty implies that person is either unemployed or employed in a relatively low wage job.

If a person in poverty is unemployed, the effect is as described in the previous paragraph. An employed person living in poverty indicates that individual receives low wages. Accordingly, a high poverty level indicates a lower mean of the overall wage distribution which results in workers lowering their reservation wages.

Measures of income levels in the respondent's area provide direct information concerning the mean of the wage distribution. Accordingly, these measures are expected to affect reservation wages in the exact opposite direction of measures of the level of poverty. Accordingly, local income levels should have a positive effect on

the wage a person expects to receive. This illustrates that an increase in income levels in the area in which a worker searches, leads to higher reservation wages for that worker.

Another factor that would affect the mean wage facing a worker is the wage that individual is currently receiving. As indicated above, individuals may search for a potentially better job while employed. Any worker who searches while employed has a readily available alternative to any job offer, their current job. Thus, a potential employer must outbid the current employer, in wage plus non-wage characteristics, to induce the worker to change jobs. As such, a higher current wage for a worker yields a higher expected wage facing that worker since the worker will not accept any job offering utility below their current wage plus non-pecuniary benefits. Accordingly, this results in a higher reservation wage.

Finally, factors such as personal characteristics might prove useful in the estimation of reservation wages. These characteristics could include gender, race, marital status, number of dependents, level of risk aversion, union status, or geographic region. These variables would serve to capture any potential systematic differences in reservation wages or labor market opportunities across different groups.

Also, the type of area in which a person lives may serve to indicate some systematic differences in reservation wages. An urban neighborhood is indicative of a higher population than a suburban neighborhood which, in turn, has a higher population than a rural area. Thus, there are more individuals in urban areas to search

for jobs than in the suburbs. However, it is also likely that there will be more jobs available in urban areas.

The result of these factors on reservation wages is unclear. However, information concerning the type of area in which a respondent resides is expected to provide information about the labor market. As such it will be used to capture systematic differences across types of areas that are otherwise unmeasurable.

A brief survey of past studies that have estimated reservation wages may show the usefulness of these variables. Table 1 highlights some important points from each of the papers discussed here. One of the earliest works to estimate reservation wages is that of Kiefer and Neumann (1979). The author's objective was to test the hypothesis that reservation wages remain constant over the duration of unemployment. The data used in this study contains information on 517 unemployed males and was collected for the Trade Adjustment Assistance.

Kiefer and Neumann used an individual's marital status and number of dependents as personal characteristics. As an indicator of cost of search the authors used the amount of unemployment benefits the individual receives. To provide information about the mean wage available to the individual, this study used age, age-squared, education, and local unemployment rate. The variables measuring unemployment benefits, age-squared, education, and unemployment rate all produced positive and significant parameter estimates. The variables for marital status and age were negative and significant in explaining reservation wages (Kiefer & Neumann, 1979).

Interestingly, the signs associated with age and age-squared indicated that age plays a non-linear role in determining reservation wages. In fact, the effect of age on reservation wage is increasing to around 26, then decreasing thereafter. The main conclusion of the authors is that reservation wages are not constant over the duration of unemployment and actually decrease by about 0.6% every week that an individual is unemployed (Kiefer & Neumann, 1979).

The next paper discussed is that of Feldstein and Poterba (1984). In this paper the authors attempted to examine how unemployed individuals react to government policy. Specifically, the impact of unemployment insurance benefits on reservation wages is explored.

To capture differences across groups of personal characteristics Feldstein and Poterba used race, gender, and a dummy variable for married males. For information concerning the net cost of search Feldstein and Poterba used variables about the level of unemployment benefits, welfare acceptance, non-working income (interest, rent, etc.), and a dummy for other workers present in the household. To show the effect of the mean of the available wage distribution, both age and education of individuals are used (Feldstein & Poterba, 1984).

Feldstein and Poterba specify several different forms of a reservation wage equation. Estimates are obtained for each of these specifications. The only variables that yielded a significant parameter estimate in a majority of the specified equations are unemployment benefits and non-working income. Both of these variables

produced positive significant parameter estimates as the above comparative statics suggest (Feldstein & Poterba, 1984).

The main conclusion of the paper is that a 10% increase in UI benefits would increase reservation wages by 4% or less. A corollary of this result is that the substitution effect of the UI benefits is larger than the corresponding income effect. That is, by increasing non-wage income, UI benefits change the price of labor thereby causing a substitution effect as well as an income effect. The positive net relationship between UI benefits and reservation wage indicates that the positive substitution effect is greater than the negative income effect (Feldstein & Poterba, 1984).

The study by Jensen and Westergård-Nielsen (1987) is considered next. This paper attempts to apply an empirical search model to the prospect of transitioning from education to the workforce. The model chosen for this work incorporates the possibility of search intensity varying by individual. The data utilized are from 306 recent lawyer's assistants graduates in Denmark surveyed during the years 1974 to 1977. Each respondent was surveyed twice, 12 months apart.

Variables intended to capture differences by personal characteristics are a dummy variable for female and area in which the respondent searched for employment, neither of which produced a significant parameter estimate. To reflect the mean wage, several variables are used. These included grades in school, age, length of work during study period (1974 - 1977), experience, part time job, and number of offers per application. To reflect the net costs of search the variables search duration, number of applications, and search intensity are used. None of the

cost variables produced a significant parameter estimate (Jensen & Westergård-Nielsen, 1987).

Number of months worked between surveys, experience, grades in school, and number of offers per application all produced positive significant estimates. Meanwhile, the relationship between reservation wages and part time job was negative and significant. There are three main conclusions reached by the authors. First, an absolute minimum reservation wage exists across individual law graduates. Second, the salary an individual receives increases with search intensity at a rapid rate. Third, the probability of receiving a job offer has a substantially positive impact on the individual's reservation wage (Jensen & Westergård-Nielsen, 1987).

The next paper discussed is by Niesing, Van Praag, and Veenman (1994). These authors attempted to estimate the likelihood of employment for both ethnic minorities and natives in the Netherlands. Also, differences in these likelihoods are addressed. The data used come from the 'Social Position and Use of Facilities by Ethnic Minorities' Survey in 1988. A subsample from the full survey consisting of the male heads of households is used for estimation purposes.

Personal characteristic variables used in this study are a set of dummy variables representing several ethnic minorities, marital status, and duration of stay in the Netherlands. To show the effect of average wage on the reservation wage the variables age, age-squared, education, and number of jobs are used. Additionally, the variables education and number of jobs are interacted with the minority dummy variables to capture any differences in the effect of these variables across different

groups. There are no variables used to reflect the cost of search on individual reservation wages (Niesing, Van Praag, & Veenman, 1994).

The relation between age and reservation wages is non-linear since the parameter estimates for age and age-squared are negative and positive, respectively. Education produces a negative significant estimate but the interaction between education and the minority dummies produce significantly positive estimates leaving a somewhat unclear picture of the effect of education on reservation wages. Similarly, number of jobs produces a negative estimate while the interaction between this variable and the minority dummy variables produce positive estimates. Being married was shown to decrease reservation wages. The authors conclude that the chances of becoming employed are substantially lower for ethnic minorities than for native Dutch. Roughly half of this difference can be attributed to differences in the individuals' characteristics. Further, the authors conclude that the remaining portion of the difference in chances of becoming employed, roughly half, was caused by employer discrimination (Niesing, Van Praag, & Veenman, 1994).

The final paper to be discussed in this brief review is by Hofler and Murphy (1994). This work attempts to estimate reservation wages using a relatively new technique, stochastic frontier regression analysis. Also, the authors test several hypotheses concerning the response of reservation wages to several exogenous factors. The data used in this study come from the January 1983 Current Population Survey. Only individuals in the survey who were employed full time during the survey week were used for estimation.

Individual personal characteristic variables that were used included dummy variables to indicate whether the individual lived inside an SMSA, lived on a farm, marital status, central city resident, female, occupation, industry, and region. Age, age-squared, tenure at current job, tenure-squared, and unemployment rate are the variables used to reflect the impact of mean wages on the reservation wage. The costs of search are reflected by the variables head of household status, number of children, family income, a dummy variable for homeowner, and the Unemployment Insurance replacement ratio (Hofler & Murphy, 1994).

Several personal characteristic variables produced significant parameter estimates. These include the dummy variables for SMSA, farm, married, central city, and female. Age is related to reservation wages in a non-linear manner with age and age-squared producing positive and negative significant estimates, respectively. Tenure has a similar relationship with tenure and tenure-squared also have positive and negative estimates, respectively (Hofler & Murphy, 1994).

The cost of search variables that result in significant parameter estimates are head of household, family income, and homeowner. Head of household and homeowner produce positive estimates while family income results in a negative significant estimate. The authors conclude that, on average, the individual's current wage exceeds their reservation wage by 25%. Also, no clear relationship is found between the UI replacement ratio and reservation wages (Hofler & Murphy, 1994).

The variables that were found to be significant in the prediction of reservation wages by the studies summarized above include the following: age, age-squared,

education, grades, experience, tenure, tenure-squared, job offers per application, part-time job, unemployment rate, number of jobs held, unemployment insurance replacement ratio, receipt of welfare, family income, head of household, homeowner, marital status, gender, and several variables indicating type of area in which a respondent resides. The significance of the personal characteristics variables cannot be explained by the above model and therefore merely serve to capture differences across the different demographic groups. The variables indicating net search costs and mean of available wages show the expected sign. Many of these variables will be used to help explain reservation wages in this study. Each of the specific variables to be used in this study will be discussed more thoroughly in Chapter 5.

This chapter has focused on developing a model where workers search for employment and wages in a world characterized by imperfect information. The model developed herein has provided a crucial step in showing the relationship between worker's preferences and occupational outcome. Specifically, this model has shown that when a worker searches over many different job categories, differences in that worker's reservation wages, d_{ijk}, are strictly caused by differences in that worker's preferences for the non-wage characteristics of those jobs. So, reservation wages are affected by a worker's tastes for the non-wage aspects of any job. This connection provides a portion of the total relationship needed. All that remains is to show how reservation wages are related to occupational outcome. This is the goal of the next chapter.

Table 1
Characteristics of Empirical Studies Estimating Reservation Wages

	Study	
Characteristic (1)	Kiefer and Neumann (1979) (2)	Feldstein and Poterba (1984) (3)
Primary Objective	Test hypothesis of constant reservation wage over unemployment duration	Examine the effect of U.I. Benefits on Reservation Wages
Data	517 males in 14 states permanently separated from their jobs in 1969-73. Data collected for the Trade Adjustment Assistance (TAA) Program.	2228 individuals who were unemployed in May 1976 and answered a supplementary questionnaire to the CPS and received U.I. benefits
Independent Variables	Educ, Dependents, Tenure(last job), Married, Urate, Age, Age^2, Educ*Age, UI Benefits, Max Duration, ln(wage last received), mean and variance of the wage distribution.	Ratio of UI to previous wage, Non-wage income ratio, Welfare acceptance, Supp UI, Other worker present, Married male, Age, White, Male, and Educ.
Significant Estimates (Sign)	Education(+), Married(-), Urate(+), Age(-), Age^2(+), UI(+), μ_w(+); Weeks unemployed(-) only in Non-constant reservation wage equation .	UI ratio(+); Non-wage income ratio(+) for job losers and others; Welfare acceptance(+) for others; Supp UI(+) for job leavers and others.
Conclusions	Job search behavior accounts for a significant portion of total unemployment. Reservation wages decrease by 0.6% per week of unemployment duration.	A 10% increase in the UI ratio causes reservation wages to increase by 4% or less.

Table 1 (Continued)
Characteristics of Empirical Studies Estimating Reservation Wages

Characteristic (1)	Study: Jensen and Westergård-Nielsen (1987) (2)	Study: Niesing, Van Praag, and Veenman (1994) (3)
Primary Objective	Apply an empirical search model, including the possibility of variable search intensity, to the transition from education to work.	Estimate likelihood of being hired for Dutch and ethnic minority groups in the Netherlands and explain differences in these likelihoods.
Data	306 recent lawyer's assistant graduates in Denmark during the period 1974 - 1977.	2098 male heads of households from the 'Social Position and Use of Facilities by Ethnic Minorities' Survey in 1988.
Independent Variables	Grades, Age, Length of work during study period, Experience, Part time job, Female, Search Area, Search duration, Number of applications, Search intensity, and Number of offers per application.	Age, Age^2, Education, Educ*Minority Dummies, Married, Duration of stay in Netherlands*Min Dummies, Number of jobs, Number of jobs*Min Dummies, and Minority Dummies.
Significant Estimates (Sign)	Number of months worked between surveys (+), Part time job (-), Experience (+), Grades (+), and Number of offers per application (+).	Age (-), Age^2 (+), Education (-), Educ*Min Dummies (+), Married (-), Number of Jobs (-), and Number of Jobs*Min Dummies (+).
Conclusions	An absolute minimum reservation wage exists for law graduates; Salary increases rapidly with search intensity; The probability of receiving an offer has a positive significant influence on reservation wages.	Employment chances are lower for ethnic minorities in the Netherlands than the native Dutch and half of the difference in employment chances is caused by employer discrimination.

Table 1 (Continued)
Characteristics of Empirical Studies Estimating Reservation Wages

Characteristic (1)	Study: Hofler and Murphy (1994) (2)
Primary Objective	Estimate reservation wages using stochastic frontier regression and test hypotheses concerning response of reservation wages to various factors.
Data	Individuals January 1983 Current Population Survey who were working full time during the survey week.
Independent Variables	Age, Age^2, Grades, Grades*Age, Tenure at current job, Tenure^2, SMSA, Farm, Married, Head of household, Central city, Unemployment rate, Female, Occupation dummies, Industry dummies, Region dummies, Number of children, Family income, Homeowner, and Unemployment Insurance replacement ratio.
Significant Estimates (Sign)	Age (+), Age^2 (-), Tenure (+), Tenure ^{2} (-), SMSA (+), Farm (-), Married (+), Central City(-), Head of household (+), Female (-), Family income (-), and Homeowner (+).
Conclusions	Current wage exceeds the reservation wage by 25%, on average, and there exists no clear relationship between UI replacement ratio and reservation wage.

Chapter 3

MATCHING

Now that the search process and the connection between tastes and reservation wages have been presented, a model for matching workers with jobs can be discussed. The objective of this model is to provide a link between reservation wages and occupation. Specifically, this model should provide for an individual's occupational outcome to be determined primarily by reservation wages. By transitivity, when this is accomplished, the result is the desired connection between worker preferences and occupational outcome. The major issue to be addressed by the matching model, developed herein, is the determination of occupational outcomes for different individuals based primarily on those individual's reservation wages.

The objective of this study is to identify and estimate segregation due to differences in tastes for job characteristics on the part of workers. Accordingly, this model of matching workers with occupations should account for, and remove, other factors affecting the job with which a worker is matched. The method chosen for this study is to create an artificial environment, a synthetic labor market, where the effect of discrimination does not enter. To create such an environment, some parameters governing the distribution of workers across occupations need to be established.

In order to understand the restrictions needed to develop the desired labor market, a discussion of a real world example of how a worker is matched with a job will be instructive. A common process for matching a worker to a job involves the

employer advertising the job opening in some way. This advertisement may, or may not, inform potential employees as to the offered wage.

Given their limited information, individuals searching for work make a decision as to whether they wish to apply for the announced opening. If the worker does apply for the job, an application is filed with the employer who collects numerous applications. The employer then examines the applications filed to create a subset of applicants who best satisfy the employer's qualifications. These qualifications can be objective or subjective in nature, or a combination of both.

When this subset of potential employees has been created, the employer then holds personal interviews in order to determine the best candidate for the job. After the personal interviews the employer then decides on a specific individual to fill the available position. The basis for choosing one individual over another can, again, be objective or subjective.

When the individual has been chosen by the employer, an offer is extended. The individual then has to decide whether or not to accept the offer. As is shown in the previous chapter, this decision is determined by the individual's reservation wage for that job relative to the offered wage. If the job is accepted, a match is made. If not accepted, the employer offers the job to another candidate and continues this process until the job is filled.

Based on this example, it is clear that the match between job and worker is affected by numerous factors including the worker's reservation wage, the wage offered, the employer's search, and the employer's preferences to name a few. While

this example is not indicative of the matching process in all cases, it does reflect the relationships between the agents involved and how the process is affected by each of the agent's preferences. Based on this real world example, we can begin to remove the determinants of the match that are not based on the worker's reservation wage.

The most obvious factor affecting the match, that is not related to the worker's reservation wage, is the preferences of the employer. These preferences are manifest in the qualifications used to determine the most qualified individuals. These preferences can include discrimination, or less onerous criteria such as previous experience. The important point is that the employer's preferences are different than the worker's and will affect the match. Because of this, the employer's preferences need to be removed from the matching process.

Another determinant of the employer-employee match, although less clear, is the employer searching for potential workers. This can influence the match in a much less subtle manner than the employer's preferences. Employer search necessarily implies a selective effort on the part of the employer to find workers. To see this consider one of the least selective search strategies, advertising a job opening in the help wanted section of a newspaper. Even though the job opening is advertised to the masses, this employer search is still selective if one or more workers do not learn of the opening. By using a search method that excludes even one potential worker, the employer can potentially affect the distribution of workers across job categories. Accordingly, employer search should be removed from the synthetic labor market being developed.

The final non-reservation wage determinant of the match highlighted by this example is the offered wage. However, unlike the previous two determinants of the match, the offered wage cannot be removed from the process. It is only when the wage offer has been received that the individual can accept or reject the offer. As demonstrated in the previous chapter, this decision is based solely on the reservation wage and the wage offered. Since the offered wage cannot be removed from the matching process, it will need to be restricted so that it is unaffected by the employer. That is, the wage offered by a given employer is assumed to be fixed regardless of the applicant.

The restrictions outlined above generate a labor market in which the employer has a very limited role. Employers do not search for employees. Nor do they determine which person they will hire based on their own preferences. Also, the employer offers a given wage regardless of the individual.

A labor market construct that satisfies these conditions is one where employers wait for applicants to arrive at their door. When an applicant arrives, the employer offers a job to the applicant at a given wage from a distribution of wages for that job. The applicant then either accepts or rejects the offer. If an offer is rejected by an applicant, there is no possibility of recall. Ex ante, workers do not know the wage an employer will offer. Also, workers are assumed to search for employment over all job categories.

By structuring the matching model in this manner, where the employer does not affect the match, the possibility of discrimination or other employer controlled

factors affecting the occupation of a worker have been removed. These assumptions concerning the potential employer, though restrictive, are sufficient to create the desired artificial environment. To see this, consider the factors affecting the match of a worker to a job within this construct.

When the worker requests and is offered employment, the employer offers a predetermined wage for a specific job. The only factor that affects the workers decision, and therefore the job match, is the worker's reservation wage for that job relative to the offered wage. Thus, three factors are shown to affect the match in this environment: the specific job offered by the employer, the wage offered, and the individual's reservation wage for that job.

Having specified the sufficient assumptions for the model, the next concern is the design of an empirical method to match a worker with a job that satisfies the above criteria. For a worker to be matched with a job in the synthetic labor market described above, two things must happen. First, the worker must receive the job offer, or, in this model, the worker must find a firm. Secondly, the worker must accept the offer after finding the employer and receiving the offer. The probability of these two events occurring will be used to match workers with specific jobs.

The probability of receiving and accepting a job offer will be referred to as the probability that worker i enters occupation j and will be written as $P(E_{ij})$. This probability can be found as the product of the probability of a worker receiving a job offer and the probability that worker accepts the job offer, given receipt of an offer. That is,

$$P(E_{ij}) = P(R_{ij} \cap A_{ij}) = P(R_{ij}) * P(A_{ij}|R_{ij}) \quad (1)$$

where $P(R_{ij})$ is the probability of the i^{th} worker receiving an offer in the j^{th} job category (this is analogous to P_j from Chapter 2) and $P(A_{ij}|R_{ij})$ is the probability of the i^{th} worker accepting employment in job j, given the offer has been received.

An important question is whether this method for matching workers satisfies the conditions above such that the match is determined by only the worker's reservation wage, the offered wage, and the type of job offered. The influence of the offered job category is captured by $P(R_{ij})$ which reflects the probability that an offered job is of type j. $P(A_{ij}|R_{ij})$ is determined solely by the relationship between the offered wage and the individual's reservation wage. Hence, this mechanism does meet the conditions above.

An additional reason for using this probability to match workers is the stochastic nature of the matching process. As discussed in the previous chapter, workers search over many different job categories and the probability of entering any of these jobs in a period is non-zero. By recognizing that matches occur with differing probabilities, this mechanism enhances the model by reflecting the randomness with which job matches occur.

Once the probabilities of a worker entering different occupations are found, that worker can be matched with the occupation where they have the highest probability of actually working. Stated differently, workers can be matched according to the rule,

$$\underset{j}{MAX}\ P(E_{ij}) \quad (2)$$

First, the $P(E_{ij})$ for a worker must be found in each occupation. Then, by comparing each of these probabilities the worker can be matched with the occupation where the probability is the greatest. This model is relatively simple but it does capture the important aspects of the matching process.

There are several different ways in which $P(E_{ij})$ can be estimated. Each of these possibilities represents a different realization of the synthetic labor market developed earlier in this chapter. The different estimations of $P(E_{ij})$ are found by estimating the probabilities of receiving and of accepting a job separately as indicated by Equation (1). Each of these probabilities will be found in two ways.

Consider first the probability that a worker receives a job offer. $P(R_{ij})$ is affected by many things, such as education, experience, or other qualifications. However, most of these determinants are eliminated with the specified assumption of total passivity of potential employers. Since employers are assumed to offer any potential employee a job when requested, employers do not affect a workers probability of receiving a job offer.

Within this construct, the probability of receiving a job offer in a particular occupation is determined by the probability of finding an employer who is hiring in that occupation. This insight will aid in the development of alternative methods of estimating $P(R_{ij})$. The simplest way to handle the probability of receiving a job offer is to assume it is constant across occupations for a particular worker. This is the first method used to determine $P(R_{ij})$. A different way of stating this assumption is that the probability of finding an employer of any specific occupation is equal to the

probability of finding an employer of any other occupation. This assumption concerning $P(R_{ij})$ provides a simplistic beginning for finding the probability of a worker entering a job.

There exists, however, at least one exogenous factor that may affect the probability of receiving an offer, even within this restrictive artificial environment. This factor is the size of different job categories. It can be argued that the larger the job category, ceteris paribus, the more likely a worker is to find an employer of that job and in turn be offered employment in that occupation. This assumes that the size of a job category is positively related to the number of employers hiring workers for that job.

$P(R_{ij})$ can be estimated in a way to capture this effect. Using the relative frequency approach, this probability can be estimated as:

$$P(R_{ij}) = \frac{\text{Number of persons employed in occupation j}}{\text{Number of persons employed in all occupations considered}}. \qquad (3)$$

This is the second method for finding $P(R_{ij})$. The denominator may need some explanation here. The reason the phrase 'all occupations considered' is used, instead of all occupations, is that not all occupations will be considered in this work due to data restrictions that will be discussed later.

Clearly this estimate of $P(R_{ij})$ is based on actual labor market data, since the relative frequency approach is used. Accordingly, this estimate of $P(R_{ij})$ is affected by factors other than taste differences between workers. However, it does reflect the

relative size of an occupation which is expected to positively affect the number of employers.

This in turn affects the likelihood of a worker finding an employer in that occupation. Necessarily then, the probability of receiving an offer is increased since employers are assumed to offer a job to any applicant. So while this estimate of $P(R_{ij})$ is not perfect, it does capture, to some extent, the effect of occupation size on the probability of an individual receiving a job offer in that occupation.

The limiting nature of the artificial environment constructed allows for only these two alternatives of estimating $P(R_{ij})$. If some of the assumptions were relaxed there would certainly be other possibilities. However, in this model we only have these two alternatives for the estimation of $P(R_{ij})$.

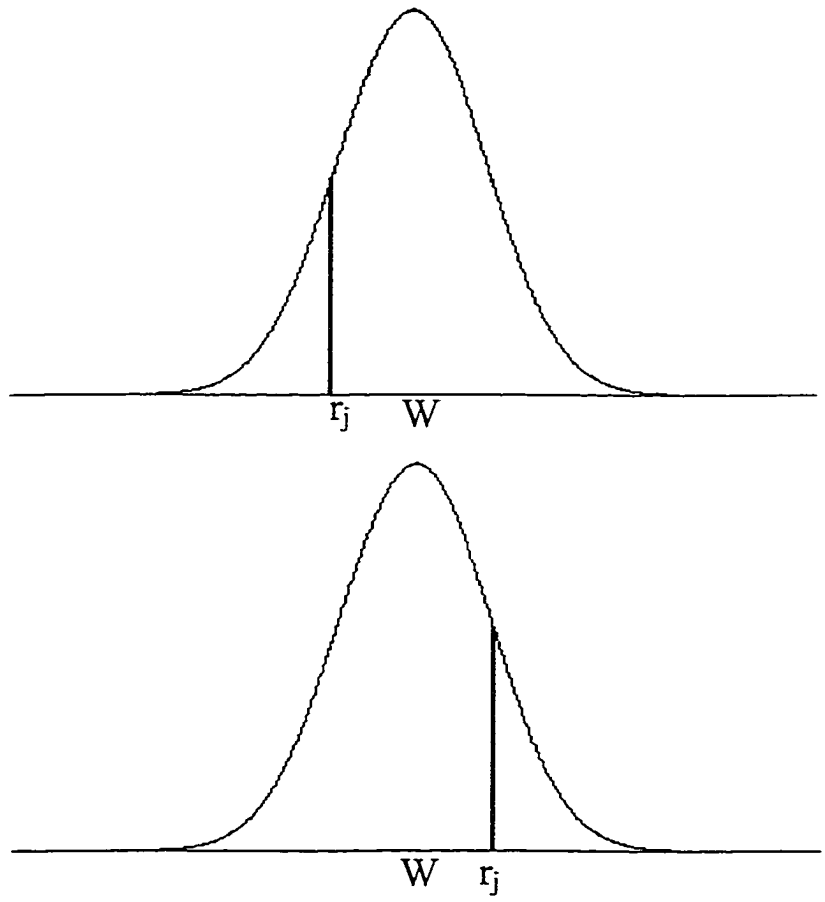
The next issue to be addressed is the probability that a worker accepts a job upon receipt of an offer. This probability, $P(A_{ij}|R_{ij})$, is a function of the reservation wage as well as the offered wage which is drawn from a distribution. The higher a workers reservation wage for a given job, the less likely they are to be offered a wage above their reservation. This, in turn, implies a smaller likelihood of the individual accepting employment in that occupation.

To see this, consider the two panels of Figure 1 below showing a probability density function of wages where r_j is the reservation wage of a worker for job j. The shaded area under the density function represents the probability that a randomly drawn wage from the distribution is greater than or equal to r_j. As such, this area represents $P(A_{ij}|R_{ij})$.

Comparing the two panels in figure 1 shows that the lower the reservation wage, relative to the wage distribution, the higher is $P(A_{ij}|R_{ij})$. If such a distribution of wages across different job categories could be found, and reservation wages had been estimated, it would be a relatively simple matter to determine each worker's $P(A_{ij}|R_{ij})$. However, before proceeding along such a path, a discussion of a possible improvement to this estimate of $P(A_{ij}|R_{ij})$ will be helpful.

Figure 1

Distribution of Wage Offers



As stated in Chapter 2, workers search for employment over many different job categories. Accordingly, the wage distributions in Figure 1 contain information about all of these jobs. The method of finding $P(A_{ij}|R_{ij})$ described above would utilize this mega-distribution of wages. However, it seems possible that within this mega-distribution there may exist some pattern in the wages for different jobs. That is, one job in which an individual searches may consistently pay lower wages than some other job where the individual also searches. Realizations of wages in low paying jobs would then tend to be clustered in the lower tail of the mega-distribution.

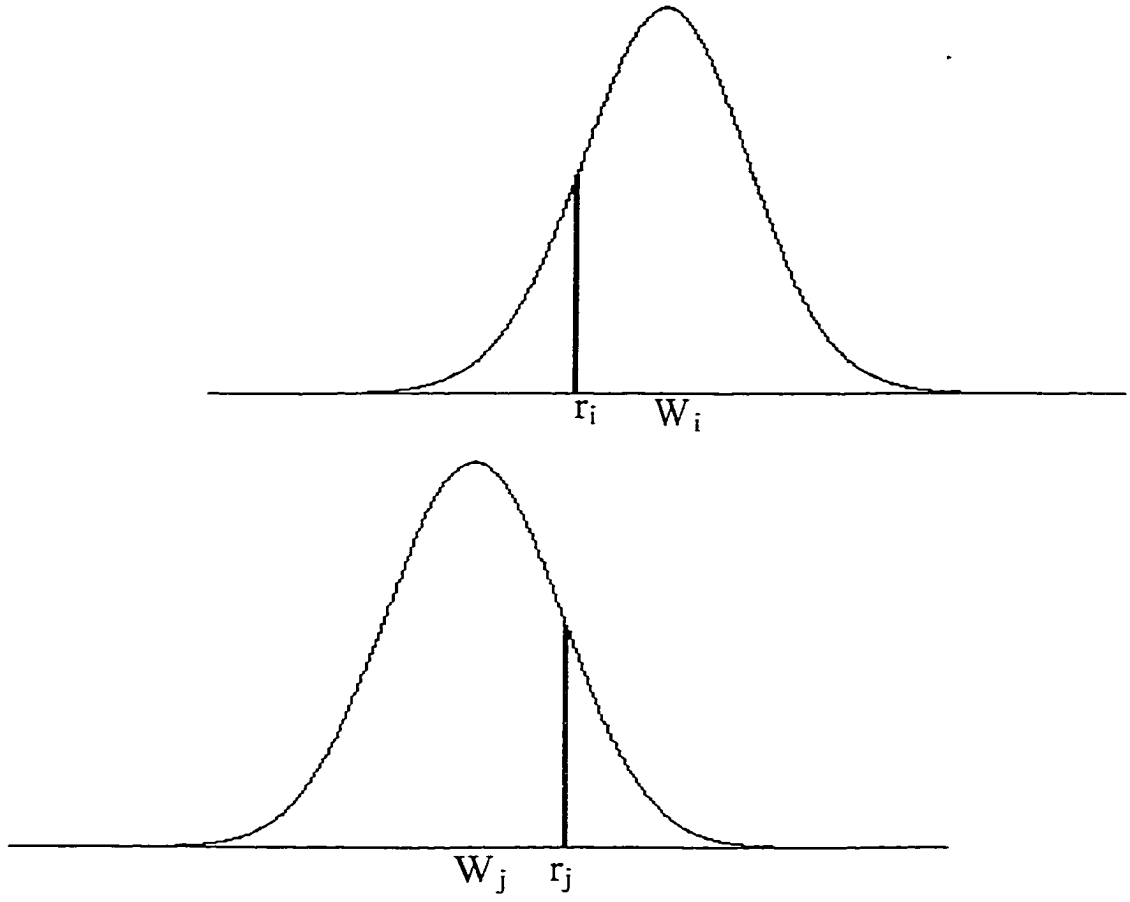
Stated differently, the mean wage in a certain job may not be represented by the mean of the mega-distribution. If this is true, using the mega-distribution of wages to estimate $P(A_{ij}|R_{ij})$ will tend to overstate the probability of acceptance for low paying jobs in which the individual searches. Conversely, this method will tend to understate an individuals probability of accepting a received offer in high paying jobs.

To see this, consider Figure 2 below. The two panels show job specific wage distributions. Clearly, the job represented in the top panel pays significantly higher wages, on average, than the job represented in the bottom panel. The individual's reservation wages represented in the figure indicate the individual finds both jobs equally distasteful. If $P(A_{ij}|R_{ij})$ were estimated using a combined wage distribution, it would predict this individual would have an equal likelihood of accepting an offered job in either of these occupations. However, Figure 2 clearly shows this is not a

realistic prediction since the probability of this worker accepting a randomly drawn wage offer is significantly higher in the first job category.

Figure 2

Distribution of Wage Offers



To account for this possibility of consistent wage differences between jobs, a different estimation method is needed. Specifically, the combined distribution of wages across all jobs cannot be used. A method of estimating $P(A_{ij}|R_{ij})$, which allows for differences in the wages of different jobs, is to compare the worker's reservation wage in a job category to a distribution of wages specific to that occupation. This is

the first method used to estimate $P(A_{ij}|R_{ij})$. It is important to note that while estimated wage distributions will be found by job category, individuals still behave as though they search over many jobs and face the mega-distribution of wages.

To accomplish the estimation of occupation-specific wage distributions, observations of wages for very precise job categories are needed. These will be obtained from the Current Population Survey (CPS). These wage observations will then be used to estimate the cumulative density function for wages, F(w), in each job category.

These distributions, however, are not perfectly suited to the needs of this work. The primary problem with these wage distributions is that they are obtained from the actual labor market and as such are affected by more than just taste differences. However, it seems a reasonable assumption that these wage distributions would change very little if all factors affecting them were removed, with the exception of taste differences between workers.

Given this assumption, these wage distributions will be used to estimate the probability of a worker accepting a randomly drawn job offer in different occupations. Using these wage distributions for specific jobs and a workers reservation wages for these jobs, it is a simple matter to find $P(A_{ij}|R_{ij})$. To accomplish this $F(w^r)$ is found by determining the proportion of wages that are below w^r. The probability $P(A_{ij}|R_{ij})$ is then equal to $1 - F(w^r)$.

Based on the different possibilities for estimating the component probabilities, there are two alternatives for estimating the probability of a worker becoming

employed in a certain job category. The first method of finding $P(E_{ij})$ will assume constant probabilities of receiving a job offer across occupations and estimate $P(A_{ij}|R_{ij})$ using the estimated wage distributions. The second method of estimating the probability of a worker entering a job category will improve on this method by estimating $P(R_{ij})$ using the relative frequency approach.

Now that the probability of a worker entering an occupation can be found, workers can be matched with the occupation where they would have the highest probability of actually working if only taste differences affected this outcome. The next relevant question is how workers are to be placed into different occupations. There are two alternative answers to this question.

The first, and most obvious, is to simply place the worker i into the occupation j where $P(E_{ij})$ is the largest. That is, for each worker place a value 1 in the occupation with the highest value for $P(E_{ij})$ and a 0 in all other occupations. This is referred to as 0/1 sorting. However, this sorting mechanism ignores part of the stochastic process of becoming employed.

To fully demonstrate this problem, consider an example. Suppose there are 100 workers to match with this model. Assume all workers are identical in exogenous factors and preferences. Each worker faces only two job prospects. The probability of any of these workers entering the first job is 0.51 and the probability for the second job is 0.49.

The expected result based on these probabilities is that 51 workers would be employed in occupation 1 and 49 in occupation 2. However, this is not the result that

the model described thus far would produce. This model, with 0/1 sorting, would place all 100 workers into the first occupation since this is the job with the largest probability of each worker being employed. Thus, sorting individuals into one specific occupation, 0/1 sorting, may distort the distribution of workers and is clearly a weakness of this model.

To address this concern a different sorting mechanism is needed. The most direct manner in which to address this concern is to place workers into occupations, by a fractional measure. That is, a worker can be placed into different occupations at the same time. This is referred to as probability sorting or p-sorting. The value, or fraction of a worker, to be placed in a given occupation, is the probability of the worker entering that job, relative to the sum of the probabilities of entering all jobs considered. This value, V_{ij}, can be written as the following:

$$V_{ij} = \frac{P(E_{ij})}{\sum_{j=1}^{n} P(E_{ij})} \tag{4}$$

Probability sorting may correct the deficiency of the 0/1 sorting process. However, a relevant concern about p-sorting is whether a distribution of probabilities indicates anything about the corresponding distribution of workers. Returning to the previous example, with p-sorting each worker is split .51/.49 into occupations 1 and 2, respectively. When these values are summed for all individuals, the result is 51 workers employed in occupation 1 and 49 in occupation 2. Thus, p-sorting produces the expected results.

By providing the relative probability of a worker entering a certain occupation, p-sorting captures the full stochastic nature of the employment matching process. At the same time it shows a distribution of probabilities of workers entering job categories. While not equivalent to a distribution of workers, it is analogous as the above example shows. So, the distribution generated by p-sorting does show how tastes affect the distribution of workers.

These two methods of placing a workers into specific occupations, combined with the two different estimates of $P(E_{ij})$, allows for four different methods by which a worker can be matched with a job. All four of these matching algorithms will be used. A summarization of the matching processes to be used may now be in order, the results of which are presented in Table 2.

First, workers will be matched to the occupation where they have the largest $P(E_{ij})$, based on estimating $P(A_{ij}|R_{ij})$ with constructed empirical distributions of wages and the individual's reservation wages. $P(R_{ij})$ will be assumed constant across different jobs. Workers will be sorted by the 0/1 process. In the second sort, workers will be placed by relative probability (p-sorting) into the occupation where $P(E_{ij})$ is largest. $P(A_{ij}|R_{ij})$ will be estimated as above and $P(R_{ij})$, again, will be assumed constant. The third method for matching workers will estimate $P(R_{ij})$ based on occupation size. $P(A_{ij}|R_{ij})$ will be estimated with reservation wages and wage distributions. Then workers will be 0/1 sorted into the job where $P(E_{ij})$ is the largest. Finally, workers will be p-sorted by $P(E_{ij})$, which will be found by estimating both $P(R_{ij})$ and $P(A_{ij}|R_{ij})$ similarly to as in the third sorting algorithm.

This chapter has developed a model for matching workers with specific occupations in a number of ways. This model has shown that a worker's occupation is dependent on reservation wages, given the assumptions within the model. Chapter 2 showed that reservation wages are driven by exogenous factors and a worker's tastes for the non-wage attributes of a job. Thus, a relationship has been shown to exist between the occupation in which a person works and that worker's preferences for the non-wage characteristics of jobs.

The primary objective of this model is to match workers to the job in which they are most likely to work if only their tastes for the characteristics of jobs affect their occupational outcome. This model is based on workers searching for employment. The search process is purely sequential which necessitates that workers form a decision rule for accepting or rejecting job offers as discussed in the previous chapter.

This decision rule is in the form of a reservation wage. To actually sort workers and create the corresponding distribution, within this synthetic labor market, both reservation wages and the distribution of wages for specific occupational categories are needed. Obtaining these estimates is the focus of the next chapter.

Table 2
Algorithms for Matching Workers with Jobs in the Synthetic Labor Market

Matching Algorithm (1)	Method of Estimating $P(R_{ij})$ (2)	Method of Estimating $P(A_{ij}\|R_{ij})$ (3)	Method of Placing Workers in Jobs (4)
1	Assumed Constant	Empirical Wage Distributions	0/1 .Sorting
2	Assumed Constant	Empirical Wage Distributions	p - Sorting
3	Relative Frequency Approach	Empirical Wage Distributions	0/1 Sorting
4	Relative Frequency Approach	Empirical Wage Distributions	p - Sorting

Chapter 4

ESTIMATION OF RESERVATION WAGES AND WAGE DISTRIBUTIONS

The method used to estimate reservation wages and wage distributions, discussed in the previous chapter, is the focus of this chapter. The data to be used for these estimations will be discussed in detail in the next chapter. This data comes from the National Longitudinal Survey, Youth Cohort (NLSY) in 1979 and the Current Population Survey (CPS) in 1979. The NLSY data will be used to estimate reservation wages while the CPS data will be used to estimate wage distributions. The estimation of reservation wages will be presented first followed by the discussion of the method for estimating the distributions of wages.

A. Estimation of Reservation Wages

Reservation wages will be estimated with data from the NLSY survey because of a series of questions that are very well suited for this task. This series of questions is the following:

> "If right now you were offered a full-time job at (hypothetical hourly wage rate) do you think you would accept it if it were (a given type of work)?" (Bureau of Labor Statistics, 1980)

If a respondent was enrolled in regular school at the time of the survey, they were instead asked:

> "If next summer you were offered a full-time job at (hypothetical hourly rate) do you think you would accept it if it were (a given type of work?" (emphasis added) (Bureau of Labor Statistics, 1980)

Each respondent was asked this question for a particular job at a wage of $2.50. They were then asked if they would accept a different job at a wage of $2.50. This question was asked for seven different job categories.

If the respondent had responded negatively to this question for any of the seven jobs, they were then asked the same question about that job at a wage of $3.50. This was repeated for all jobs in which the respondent indicated they would not accept employment at an hourly wage of $2.50. For any job that a respondent indicated they would not be willing to accept employment at a wage of $3.50 they were asked the question a final time at a wage of $5.00.

The following are the job categories for which these questions were asked: (a) washing dishes, (b) working in a factory, (c) working as a cleaning person, (d) working at a check-out counter in a supermarket, (e) working cleaning up neighborhoods, (f) working at a hamburger place, and (g) working away from home in a national park. The job category "working in a factory" is not used hereafter since different respondents may have very different ideas about the characteristics of this job, thereby yielding inconsistent information about reservation wages and tastes. Also, the job categories "working cleaning up neighborhoods" and "working away from home in a national park" are not used because there is no close match to these jobs in the Current Population Survey data which is used to make comparisons of reservation wages to the distribution of actual wages.

This series of questions allows for the identification of the range of an individual's reservation wage in these job categories. The ranges of reservation

wages are (i) $2.50 or less, (ii) $3.50 or less but greater than $2.50, (iii) $5.00 or less but greater than $3.50 and, (iv) greater than $5.00. Using this information, as well as demographic and individual characteristics identified by the survey data, continuous value reservation wages can be estimated.

A logical question at this point is why should continuous value reservation wages be estimated. First, it should be clear from Chapter 3 that specific values, instead of a range of values, for reservation wages are needed in order to match workers with occupations. It would not be possible to compare a range of reservation wage values to a distribution of wages and generate the probability of a particular individual being matched to a certain occupation.

The next question is why reservation wages are estimated as opposed to some other method of arriving at a specific value. To answer this question, consider the alternatives. Since the data does not allow for direct measurement of a reservation wage value, the only possible techniques available for arriving at a specific value are estimation and assumption.

There are numerous methods of assuming a reservation wage for an individual. One of the easiest of which is to simply use the lower bound of the category as the reservation wage. If the respondent identified category (iii) as the range of their reservation wage, then, by assumption, their reservation wage would be $3.50. However, several major weaknesses with this approach are easily seen.

First, the individual's reservation wage in this job may actually be much closer to $5.00 than $3.50. Assuming the reservation wage to be the lower bound of

the category thus provides very misleading information about this individual as well as their preferences concerning this occupation. Second, how does one handle a situation where the individual identifies category (i) as containing their reservation wage for a certain job. Obviously, assuming a reservation wage of $0.00 is not a prudent choice. The question then becomes "How low should it be?" There is no good answer to this question since it would merely be an assumed value and would be based on no other information than knowing that the reservation wage is no more than $2.50.

Third, and most importantly, since several different job categories are being compared, it is quite possible that a respondent may identify one wage category as containing all of their reservation wages. If this happens, merely assuming the reservation is at the lower bound would imply that the individual has identical preferences for all four job categories. It is highly unlikely that an individual would find working at a checkout counter exactly as distasteful as working as a cleaning person or washing dishes. It is a straightforward assertion that simply assuming the upper bound of the reservation wage category would be plagued by the same problems as assuming the lower bound.

Another method of assuming a specific value for reservation wages is to assume the individual's reservation wage in an occupation is the arithmetic mean of the upper and lower bound of the identified range containing the reservation wage. That is, for a reservation wage identified as being in category (iii), the assumed value

would be $4.25. Clearly, this approach does not overcome the major problems with the prior examples.

In order to proceed we need specific information about each reservation wage for every respondent since this is how the individual's tastes are quantified. Without this information about the individual's reservation wage, individual workers could not be matched with jobs. If workers cannot be matched to an occupation individually, a distribution of workers, across jobs, cannot be constructed. Since a distribution of workers across occupations is required to determine the amount of gender segregation that is caused by self selection, assuming a specific value for reservation wages will not suffice. The only alternative technique is to estimate reservation wages to which the discussion now turns.

As indicated in Chapter 2, reservation wages are assumed to be a function of individual characteristics and demographic variables[4]. That is, assume reservation wages are determined by the following function:

$$\ln(r_{ij}) = X_i\beta_j + \sigma_j\varepsilon_{ij} \qquad (1)$$

where r_{ij} is the i^{th} individual's reservation wage for the j^{th} job, X_i is the vector of individual i's characteristic and demographic variables, β_j is the corresponding parameter vector related to the j^{th} job category. Additionally, ε_{ij} is a logistically distributed error term with a mean of zero and variance of $\pi^2/3$, and σ_j is the positive scale parameter associated with the error term. A logical question here is why the

error term in this equation is assumed to be logistically distributed. This matter will be fully addressed later in the chapter.

Obviously, estimation of equation (1) can not simply proceed by a Least Squares method since the dependent variable contains only limited information. The information is limited in that the dependent variable is categorical. To estimate reservation wages a categorical dependent variable, C_{ij}, representing the self reported category of the i^{th} individual's reservation wage for the j^{th} job category is defined as follows:

$$C_{ij} = \begin{cases} 1, \text{ if } \ln r_{ij} \leq \ln(\$2.50) \\ 2, \text{ if } \ln(\$2.50) < \ln r_{ij} \leq \ln(\$3.50) \\ 3, \text{ if } \ln(\$3.50) < \ln r_{ij} \leq \ln(\$5.00) \\ 4, \text{ if } \ln(\$5.00) < \ln r_{ij}. \end{cases} \tag{2}$$

Estimation of β_j and σ_j can be accomplished by an iterative maximum likelihood process based on the reported category of the individual's reservation wage and that individual's characteristics (C_{ij}, X_i). In this case, the log likelihood function for the j^{th} occupation is of the following form (Maddala, 1983):

$$\ln L_{ij} = \sum_{i=1}^{n} \sum_{k=1}^{4} \delta_{ijk} \ln \left[F\left(\frac{u_k - X_i\beta_j}{\sigma_j} \right) - F\left(\frac{l_k - X_i\beta_j}{\sigma_j} \right) \right] \tag{3}$$

where δ_{ijk} is a dummy variable which takes the value 1 if the i^{th} respondents' reservation wage for job j is identified as being in category k. That is, if the first respondent's reported reservation wage category for the job of "washing dishes" (the

---

[4] A discussion of what variables are considered to affect an individual's reservation wages are

first job category) was greater than $5.00 per hour (the fourth wage category), then $\delta_{114} = 1$ and $\delta_{111} = \delta_{112} = \delta_{113} = 0$. F(•) represents the cumulative density function for the logistic distribution, which is

$$F(x) = \frac{1}{1 + \exp(-x)}. \quad (4)$$

Also, u_k is the ln(upper bound of the identified category of the reservation wage), and l_k is the ln(lower bound of the identified category of the reservation wage). The upper bound of the fourth category is assumed to be 15 and the lower bound of the first category is assumed to be 0.1. That is, reservation wages are assumed to line between $0.10 per our and $15 per hour. This assumption is made to avoid computing errors, however, no adjustment is made to the likelihood function. This is due to the fact that truncating the distribution at $0.10 and $15 merely adds a constant term to the log-likelihood function which can be ignored.

This log-likelihood function represents the sum, across individuals, of the likelihood that the i^{th} individual's reservation wage for the specified j^{th} job category is in the reported wage category based on that individual's characteristics. To see this, examine each of the individual terms in equation 3. The fractions in the log-likelihood function provide a number similar to a Z-value for the normal distribution such that the distribution is assumed to be centered around $X_i\beta_j$ with a standard deviation of σ_j. Taking F(•) of this value provides the probability of any randomly drawn observation from the logistic distribution taking a value less than the fraction.

---

presented in Chapter 5.

Subtracting the cumulative probability for the lower number from the cumulative probability for the higher number provides the probability that a randomly drawn observation from the logistic distribution is between the two numbers. That is, the calculation provides the likelihood that the individual's reservation wage is within the identified range based on the individual's characteristics. Summing these likelihoods for all individuals provides the log-likelihood function for occupation j. Maximizing this log-likelihood function provides estimates of β_j and σ_j.

This simple estimation of equation 3, however, ignores potentially useful information. For example, it is quite likely that a certain individual will report similar ranges of reservation wages for different jobs. That is, if a respondent has a higher than average reservation wage for job j, it is likely that respondent will have a higher than average reservation wage for job k. This implies that ε_{ij} and ε_{ik} are dependent or $Cov(\varepsilon_{ij}, \varepsilon_{ik}) \neq 0$. Therefore, C_{ik} contains information that will improve the estimation of β_j and σ_j and hence $\ln(r_{ij})$.

To incorporate this information, a dummy variable approach is taken using δ_{ijk}[5]. Dummy variables representing the identified reservation wage category for the other jobs will be added to the equation estimating reservation wages for the current job. To construct the dummy variables, the i^{th} individual's vector of the reported

[5]As is discussed by McCue & Reed (1996) the dummy variable approach is a second best alternative; however, it is more reliable and easier to use than the first choice of specifying a joint distribution of all four error terms.

category of reservation wage for the second through fourth jobs, to be used in the estimation for the first job category, can be specified as

$$\Delta_{i1} = (\delta_{i22},\delta_{i23},\delta_{i24},\delta_{i32},\delta_{i33},\delta_{i34},\delta_{i42},\delta_{i43},\delta_{i44}) \tag{5}$$

where, as above, $\delta_{ijk} = 1$ if $C_{ij} = k$, 0 otherwise for k = 1,2,3,4. Again, the variable δ_{ijk} indicates the range of the i^{th} respondent's self-reported reservation wage for occupation j. The omitted wage category is $\ln(r_{ij}) \leq \ln(\$2.50)$ or $C_{ij} = 1$ or $\delta_{ij1} = 1$. To demonstrate how this variable is constructed assume that the first individual reports $C_{11} = 2$, $C_{12} = 3$, $C_{13} = 4$, and $C_{14} = 1$, then

$$\Delta_{11} = (0, 1, 0, 0, 0, 1, 0, 0, 0),$$

$$\Delta_{12} = (1, 0, 0, 0, 0, 1, 0, 0, 0),$$

$$\Delta_{13} = (1, 0, 0, 0, 1, 0, 0, 0, 0),$$

and

$$\Delta_{14} = (1, 0, 0, 0, 1, 0, 0, 0, 1).$$

Using this additional information the full model is then

$$\ln(r_{ij}) = X_i\beta_j + \Delta_{ij}\gamma_j + \sigma_j\varepsilon_{ij}. \tag{6}$$

The log-likelihood function for job category j is then

$$\ln L_j = \sum_{i=1}^{n}\sum_{k=1}^{4}\delta_{ijk}\ln\left[F\left(\frac{u_k - (X_i\beta_j + \Delta_{ij}\gamma_j)}{\sigma_j}\right) - F\left(\frac{l_k - (X_i\beta_j + \Delta_{ij}\gamma_j)}{\sigma_j}\right)\right]. \tag{7}$$

This estimation produces estimates of β_j, γ_j, and σ_j.

Once these estimates have been obtained, a prediction of the natural log of reservation wages, $\hat{p}_{ij}$, can be found as $\hat{p}_{ij} = X_i\hat{\beta}_j + \Delta_{ij}\hat{\gamma}_j$. It should be noted, however, that $\hat{p}_{ij}$ is not the optimal estimator of $\ln(r_{ij})$ within this framework. This is

because $\hat{p}_{ij}$ may lie outside the reported range of the reservation wage for job j by individual i. For example, a respondent may indicate that $C_{i3} = 2$ [ln(\$2.50) $\leq$ ln(r_{ij}) < ln(\$3.50)], yet $\hat{p}_{i3} = 1.504$ (approximately ln(\$4.50)).

In this case $\hat{p}_{i3}$ directly contradicts C_{i3} as reported by the individual. As it turns out, this is a relatively frequent occurrence. Accordingly, this potential contradiction between what the individual reports and what is estimated needs to be addressed. The reported category of the reservation wage, C_{ij}, can be used to address this concern and improve the estimation of reservation wages.

In most models little can be said about the error term other than it's expected value is zero and variance is assumed finite and constant. However, in this model, there is information available concerning the range of each individual error term. The estimate $\hat{p}_{ij}$ together with C_{ij} indicate something about the magnitude, as well as the sign, of the error term.

Suppose that a certain respondent has a $\hat{p}_{ij}$ = 0.91629 (approximately ln(\$2.50)) for the job "washing dishes", and reported a reservation wage of more than \$3.50 but less than or equal to \$5.00 for this job. The information available indicates that the error term for this observation is most likely positive. Also, $\hat{p}_{ij}$ and C_{ij} suggest the error is likely between ln(\$1.40) = [ln(\$3.50) - ln(\$2.50)] and ln(\$2.00) = [ln(\$5.00) - ln(\$2.50)]. That is, $\sigma_1\varepsilon_{i1} \in$ (ln(\$1.40), ln(\$2.00)]. This example shows how $\hat{p}_{ij}$ and C_{ij} can provide information about ε_{ij}.

Using the available information, $\sigma_j e_{ij}$ can be approximated by the following:

$$E(\sigma_j\varepsilon_{ij} \mid \hat{p}_{ij},\ C_{ij},\ \hat{\sigma}_j) = \hat{\sigma}_j \int_{L(\hat{p}_{ij})}^{U(\hat{p}_{ij})} \varepsilon \frac{f(\varepsilon)}{F(U(\hat{p}_{ij})) - F(L(\hat{p}_{ij}))} d\varepsilon \qquad (8)$$

where $F(\bullet)$ is the cumulative density function for the logistic distribution. Also, $f(\bullet)$ is the probability density function for the logistic distribution which is

$$f(x) = \frac{\exp(x)}{(1 + \exp(x))^2}. \qquad (9)$$

Also, $U(\hat{p}_{ij}) = [(\ln(\text{upper bound of the reported category of the reservation wage}) - (\hat{p}_{ij}))/(\hat{\sigma}_j)]$ and $L(\hat{p}_{ij}) = [(\ln(\text{lower bound of the reported category of the reservation wage}) - (\hat{p}_{ij}))/(\hat{\sigma}_j)]$. That is, $U(\hat{p}_{ij}) = [(u_k - \hat{p}_{ij})/(\hat{\sigma}_j)]$ and $L(\hat{p}_{ij}) = [(l_k - \hat{p}_{ij})/(\hat{\sigma}_j)]$.

The essence of this technique is to approximate $\sigma_j e_{ij}$, given C_{ij}, $\hat{p}_{ij}$, and $\hat{\sigma}_j$, such that the result is a final estimate of the reservation wage within the reported category. Using $U(\hat{p}_{ij})$ and $L(\hat{p}_{ij})$, as defined above, as the bounds of integration specifies the region of the distribution of ε_{ij} that corresponds to an error term explaining the difference between $\hat{p}_{ij}$ and C_{ij}. Within the specified region, the integration finds the expected value of ε_{ij} by taking each possible value of ε_{ij} multiplied by the probability density function of the logistic distribution at that value.

The division by the total area under the logistic distribution in the region specified normalizes the expected value so that the result is a number within the specified interval. The entire process is identical to finding a weighted mean where each possible ε_{ij} in the specified range is an observation, the height of the logistic density function at each ε_{ij} is the weight of each observation, and the cumulative density of the logistic distribution in the region specified is the sum of all weights.

Thus, this approximation technique yields a value for $\sigma_j\varepsilon_{ij}$ that incorporates the reservation wage category reported by the respondent.

The above approximation technique used to find $E(\sigma_j\varepsilon_{ij} \mid \hat{p}_{ij}, C_{ij}, \hat{\sigma}_j)$ shows why the error term is assumed to be logistically, instead of normally, distributed. $E(\sigma_j\varepsilon_{ij} \mid \hat{p}_{ij}, C_{ij}, \hat{\sigma}_j)$ is found by integrating the probability density function of ε. If the error term had been assumed to be normally distributed, this approximation would not be possible.

This is because a definite integral of the normal distribution does not exist in a closed form solution (Judge et al., 1988). A normally distributed error term would not allow for ε to be approximated in this manner and valuable information would be lost. Thus, assuming the error term to be logistically distributed is the next logical choice since it is similar to the normal distribution and it's probability density function does exist in closed form.

The final estimate of a respondent's reservation wage for a particular job, using all available information, is then:

$$\ln(\hat{r}_{ij}) = X_i\hat{\beta}_j + \Delta_{ij}\hat{\gamma}_j + E(\sigma_j\varepsilon_{ij} \mid \hat{p}_{ij}, C_{ij}, \hat{\sigma}_j). \qquad (10)$$

This technique ensures that the final predicted $\ln(\hat{r}_{ij})$ is within the same interval as the reported reservation wage, as previously stated. Results of the estimation of reservation wages are presented in Chapter 6 along with other sets of results.

B. Estimation of Wage Distributions

Having discussed the method of estimation of reservation wages, the discussion now turns its focus to the method used to estimate the relevant wage distributions. There are at least four different wage distributions to be estimated, one for each job category considered in this work. As previously stated, CPS data will be used for this estimation. From this data, hourly wages are constructed for each individual employed in one of the four occupations considered herein. The three digit CPS occupation codes associated with each of the four occupations are: (i) washing dishes - "dishwasher" (913); (ii) working as a cleaning person - "chambermaids and maids, excluding private household" (901), "cleaners and charwomen" (902), "janitors and sextons" (903), and "maids and servants, private household" (984); (iii) working at a check-out counter in a supermarket - "cashier" (310); and (iv) working at a hamburger place - "food counter and fountain workers" (914).

The cumulative distributions of the natural log of hourly wages for these occupations are shown in Figures 3 through 6 at the end of this chapter. Using only one wage distribution for each job category, however, would ignore potentially significant variations in the wage distributions faced by different individuals. One major cause of such variation could be the geographic location of a worker. It is quite likely that a worker in Kansas City, Missouri would not face the same distribution of wages as a person in New York, New York.

What is needed to compensate for the possible geographic differences in wages is some form of a distribution of wages for each area in which NLSY survey

respondents are located. The specific indicator of an individual's geographic area that is used for this work is the Standard Metropolitan Statistical Area (SMSA). The SMSA is chosen as the indicator of a geographic area because it is the smallest geographic area reported in both the CPS and NLSY.

The state of residence could be used as the geographic area indicator but this could mask some of the wage variation across location. It seems a plausible assertion that the wage structure within a given SMSA would be more homogeneous than it would be across an entire state. One problem with using SMSA as the geographic indicator is that the CPS data in 1979 restricts its attention to only forty-four SMSAs and, as such, some observations from the NLSY cannot be used.

A distribution of wages for each of the four job categories in all forty-four SMSAs considered by the CPS data is needed. Four job categories multiplied by 44 specific SMSAs results in a total of 176 wage distributions. The easiest way to obtain the 176 needed distributions is to simply separate the CPS data by occupation and SMSA, then construct the appropriate location-job specific wage distributions.

This, however, is not a feasible solution to the problem because there are some SMSAs in which there are no observations of wages for certain job categories. Also, there are numerous instances where there are very few observations of wages in specific SMSA-occupation pairs. Since the desired wage distributions cannot be estimated in this direct manner, a different estimation technique must be found.

To construct all 176 wage distributions needed, two simplifying assumptions are necessary because of data limitations. First, wages across occupations are

assumed to be affected in a similar manner by the geographic location. That is, if the average wage for one occupation is lower in a particular SMSA, then the average wage for all occupations is assumed to be lower in that SMSA.

Secondly, the distribution of wages for a specific occupation will be assumed to have a similar shape and level of dispersion in every SMSA. This assumption allows for the use of the existing distributions of wages for each of the four job categories. All that is needed is some way to shift each distribution, affecting only the mean, according to differences in wages across SMSAs.

Since the existing distributions of wages are to be shifted, a measure of the magnitude and direction of the required shift of the wage distributions, for each SMSA, is needed. To find an appropriate measure, a determination of the relationship between wages and each SMSA is required. One possibility for quantifying this relation is the simple correlation coefficient. This would indicate the direction of the relation, but it would not provide a specific magnitude. So, the correlation coefficient does not provide the necessary information.

Another possibility for quantifying the relationship between wage and SMSA is with regression analysis. Wages could be regressed on a series of dummy variables representing SMSA and occupational category. The coefficient estimates associated with the SMSA dummy variables would provide a simple measure of the relationship between wages and geographic location. Most importantly, the parameter estimates would indicate both direction and magnitude of the relationship. Since this approach

provides both of the characteristics needed for the relation of wages to a specific SMSA it will be used here.

Let a vector of dummy variables identifying an individual's unique SMSA and occupation be written as:

$$Z_i = \{s_1, s_2, \ldots\ldots, s_{42}, s_{43}, D, C, S\} \tag{11}$$

where $s_j = 1$ if respondent resides in the j^{th} SMSA, 0 otherwise. The terms D, C, and S are dummy variables indicating the occupation of the individual where the letter corresponds to the first letter of the job categories. For example, D = 1 if the respondent was employed washing dishes, 0 otherwise. The forty-fourth SMSA and the job category of "working at a hamburger place" (BURGER) are omitted from the vector of dummy variables.

Then the regression of interest estimates the following function:

$$\ln(w_i) = \alpha + Z_i\phi + \eta_i \tag{12}$$

where w_i represents an individual's calculated hourly wage and η_i is a white noise error term. Estimates of α and ϕ can be found by the Ordinary Least Squares method since η_i is assumed to be a well behaved error term. The parameter estimates within $\hat{\phi}$ associated with the dummy variables representing SMSAs can then be used to shift the distribution of wages for each job category. The result is an estimation of the distribution of wages faced by workers in all 44 SMSAs for each of the four different job categories. Each of the resulting 176 distributions reflect the effect of both location and occupation on wages.

This chapter has presented the method of estimation used to find individual's reservation wages as well as the wage distributions faced by these individuals. The estimation of reservation wages relies on a series of question from the NLSY. These questions allow for identification of ranges of an individual's reservation wages for several different occupations.

Using this categorical data, the reservation wages can be estimated as a continuous variable with maximum likelihood estimation. There are two distinct improvements that can be made to the estimated reservation wages by incorporating additional information. Using all available information, an estimate of an individual's reservation wage is found that corresponds to the individual's survey response.

The estimation of the relevant wage distributions is accomplished by using CPS data. By constructing wages for individuals in the different job categories, a wage distribution for each occupation is found. These distributions are then shifted to account for differences in wages in different geographic locations. The result is an estimated distribution of wages for each occupation that is specific to each individual's SMSA.

The next chapter will focus on the data used in the estimation of both reservation wages and wage distributions. Specifically, Chapter 5 will discuss the nature of the National Longitudinal Survey, Youth Cohort and the Current Population Survey. The discussion will also present, in detail, the variables used to estimate reservation wages as well as the wage distributions.

Figure 3: Cumulative Density Function of ln(Hourly Wage) for Occupation "Burgers"

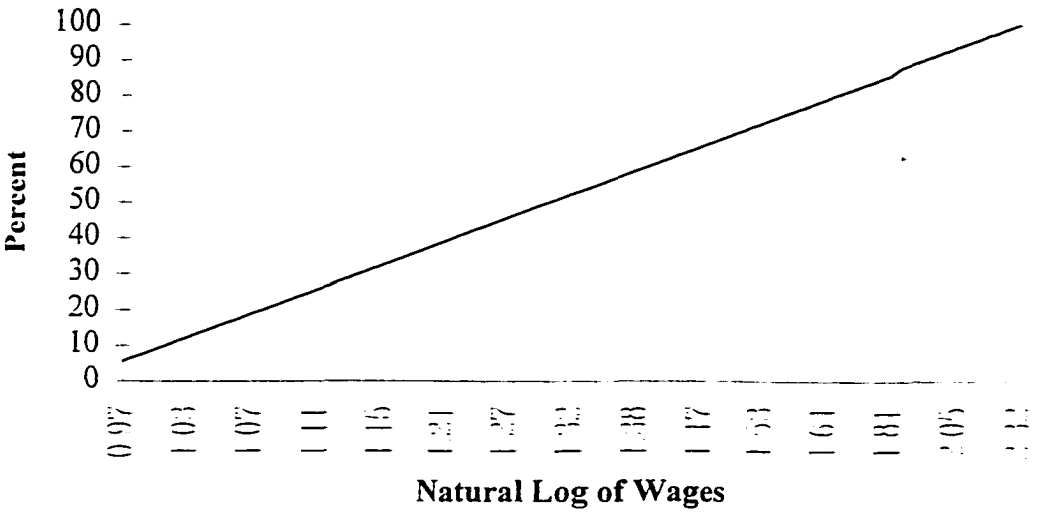


Figure 4: Cumulative Density Function of ln(Hourly Wage) for Occupation "Cleaning"

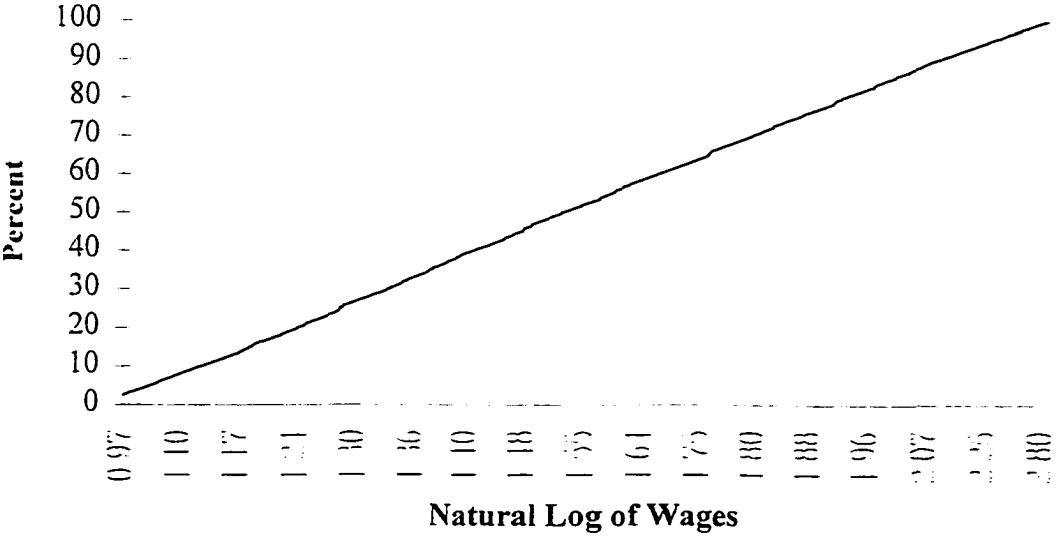


Figure 5: Cumulative Density Function of ln(Hourly Wage) for Occupation "Dishes"

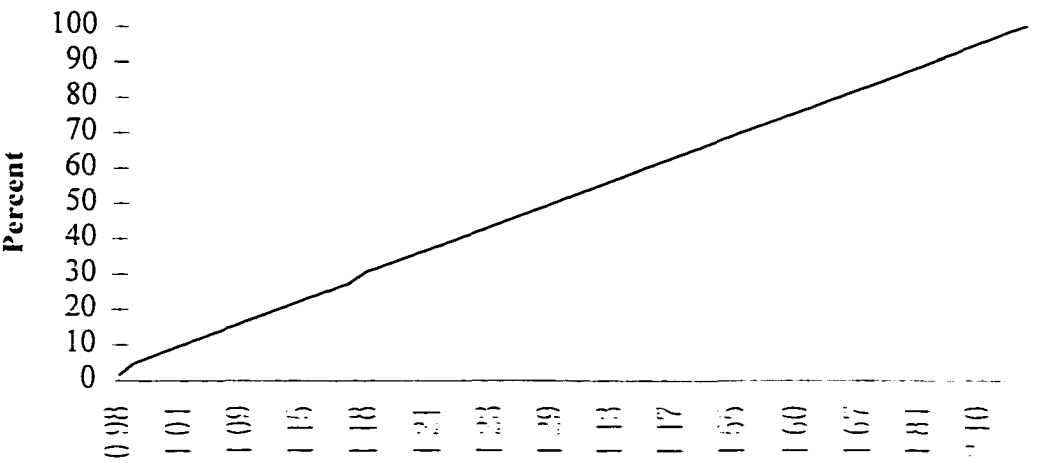
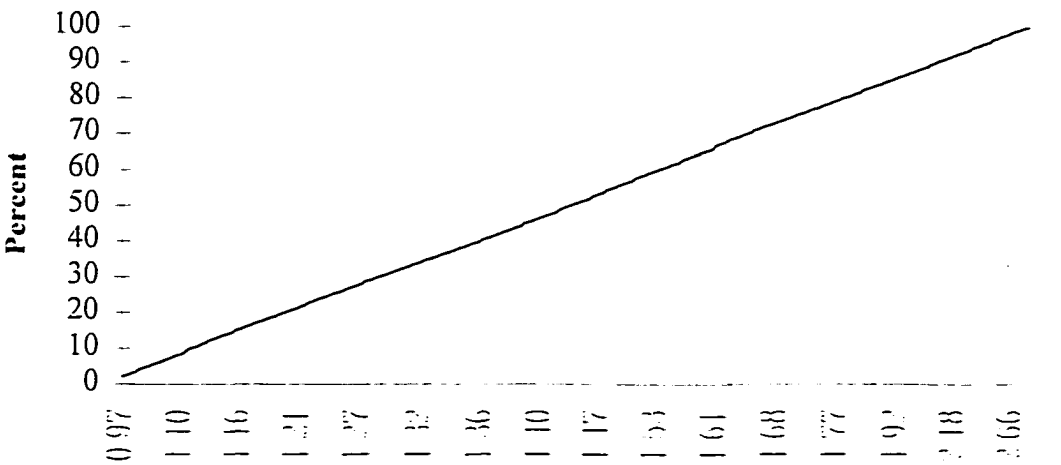


Figure 6: Cumulative Density Function of ln(Hourly Wage) for Occupation "Supermarket"



Chapter 5

DATA

The only steps that remain in the process of examining the role of taste differences in occupational gender segregation is to carry out the necessary estimations, match workers with occupations, and examine the predicted levels of segregation. However, before proceeding with these final steps, the data used for estimation warrants discussion. The next chapter will then address the results of this work. As the previous chapter indicated, two data sets are used in this study, the National Longitudinal Survey, Youth Cohort (NLSY) and the Current Population Survey, March Annual Demographic File (CPS), both in 1979. The data used to estimate reservation wages, the NLSY, will be discussed first.

The National Longitudinal Survey, Youth Cohort is an annual survey initiated in 1979 by the Bureau of Labor Statistics and the Center for Human Resource Research of The Ohio State University. The NLSY is an extension of the original National Longitudinal Surveys (NLS) which originated in 1966 and are comprised of four cohorts: mature men, mature women, young men, and young women. The NLSY was added to the NLS in order to replicate the surveys of the original young cohorts and allow for comparisons between the surveys. One specific reason the NLSY was added to the NLS was to allow for the evaluation of the 1977 amendments to the Comprehensive Employment and Training Act which provided more government sponsored training and employment.

The NLSY is comprised of three distinct sub-samples. The first represents a cross-section of the population of the United States that was born during the years 1957 through 1964. That is, respondents were between age 14 and 21 on January 1, 1979, the original survey year. The second sub-sample, referred to as the supplemental sample, within the NLSY is a group intended to over-represent blacks, Hispanics, and underprivileged whites, again all between ages 14 and 21 at the beginning of 1979.

The third group within the NLSY, is a sample of individuals serving in the military at the time of the survey. All the individuals in this third group were between the ages of 17 and 21 as of January 1, 1979. The total number of survey respondents for the NLSY in the original survey was 12,686. The cross-sectional sample contained 6,111 of the total respondents while the supplemental sample consisted of 5,295 respondents and the military sample was made up of 1,280 individuals.

The full NLSY can be broken down into the following groups by individual characteristics. The NLSY consisted of 6,403 males and 6,283 females. The cross-sectional sample was comprised of 3,003 males and 3,108 females. The supplemental sample contained 2,576 males and 2,719 females. The military sample was made up of 824 males and 456 females.

The breakdown by race for the entire NLSY shows that 7,510 individuals were white, 3,174 were black, and 2,002 were Hispanic. The cross-sectional sub-sample of the NLSY consisted of 4,916 white respondents, 751 black respondents,

and 444 Hispanics. Within the supplemental sample 1,643 individuals were identified as poor white, 2,172 respondents were black, and 1,480 were Hispanic.

The interviews for the 1979 NLSY were conducted in person by interviewers from the National Opinion Research Center at the University of Chicago between late January and the middle of August. This is one of the reasons that this particular data set is used to estimate reservation wages. Since all of the information is collected in a relatively short time frame, there is no need to be concerned with differences over time.

This, however, is not the primary reason for the use of NLSY data here. Most importantly, this data allows for the estimation of reservation wages in multiple job categories for a given individual. This is critical to this work. Without reservation wages in several job categories for each worker, differences in worker's preferences for the non-wage characteristics of the jobs could not be quantified. Consequently, matches between worker and job could not be made on the basis of these worker preferences. The NLSY is the only known data set that satisfies this need for individual specific reservation wages across jobs which is why it is used to estimate reservation wages.

The variables collected for each individual in the NLSY can be separated into three major categories: labor market experience variables, human capital and other socioeconomic variables, and environmental variables. The dependent variables, ranges of reservation wages, are derived from the series of questions described in

Chapter 4. This previous discussion has provided sufficient detail concerning the dependent variable and, as such, no further presentation is warranted.

As was discussed in Chapter 2, independent variables that provide information about, and can be used to estimate, reservation wages, are any exogenous variables affecting the net cost of search, mean of the available wage distribution, and personal characteristics. The discussion will begin by addressing variables that reflect the mean of the available wage distribution. Variables affecting the net cost of search will then be presented and the discussion will conclude by presenting personal characteristics. Table 1 lists all variables used to estimate reservation wages as well as their means for two sub-groups of the sample, males and females. Throughout the discussion of the independent variables the variable name will be indicated as (VARNAME).

Factors that are expected to reflect the mean wage available to an individual can be broken into two primary groups, human capital factors and labor market factors. Some of the factors in these groups have been briefly discussed in Chapter 2. A more detailed analysis will be presented here. The discussion will begin with human capital factors and then cover labor market factors.

There are numerous variables that may indicate the level of human capital an individual has accumulated. The first such variable is a person's age (AGE). In this context, AGE acts as a proxy since it is not a direct measure of human capital. It is expected that an older individual will have more human capital than a younger individual since the older person has had more time to accumulate education or

experience. Since AGE is a proxy for human capital, it is expected to be positively related to the reservation wage.

As well as age, other variables reflecting the amount of human capital an individual has accumulated are used to help estimate reservation wages. Unlike age, however, these other variables are more direct measures of human capital. The first variable in this class is the highest grade of school a respondent has completed (GRADE). This variable takes on integer values between 0 and 18. Values beyond 12 represent college level education. Each extra year of education is expected to provide an individual with additional human capital. As such, GRADE is expected to positively affect reservation wages.

A factor that might affect the amount of human capital an individual accumulates is the cost of such accumulation. A variable that may capture differences between individuals costs of attaining human capital is the education level a survey respondent expects to complete (EDUCGOAL). An expectation of completing a higher number of years of education indicates a lower self-assessment of that individual's cost of human capital attainment.

Numerous factors affect an individual's cost of human capital investment. One such factor is the individual's endowment of intelligence. A smart person need not study as hard as a dullard and, accordingly, has a lower cost to any given educational investment. Thus, a lower self-assessed cost of human capital investment might indicate a higher level of intelligence which implies a higher current level of

human capital. Accordingly, the variable EDUCGOAL is expected to positively affect reservation wages.

Another set of variables directly reflecting human capital levels are specifically related to a respondent's high school education. The first variable in this group is a dummy variable indicating whether the individual is a high school graduate (HSGRAD). That is, this variable takes a value 1 if a respondent had completed 12 years or more of regular schooling, 0 otherwise.

This variable, while derived from GRADE, captures an influence on reservation wages that may otherwise be missed. Ignoring whether a person graduated from high school would implicitly assume that the 12th year of education is just as valuable as the 11th year. This would imply that a high school diploma has no value in human capital terms beyond the years of education it imparts. Being a high school graduate is expected to increase human capital and therefore, reservation wages.

Another high school factor that is of interest is the type of program a person studied. Several types of high school programs are identified in the NLSY. Two dummy variables are constructed for high school program type.

The first takes a value 1 if the program was identified as primarily vocational or commercial, 0 otherwise (HSVOCCOM). The second dummy variable takes a value 1 if the respondent's high school program was identified as primarily general, 0 otherwise (HSGENPRG). The omitted category is a college preparatory high school program.

Similar to EDUCGOAL, these high school type variables may indicate the individual's cost of human capital attainment beyond high school. Individual's with lower expected costs are more likely to attend college and therefore pursue a college preparatory high school program. Since HSVOCCOM and HSGENPRG both indicate that the individual is not in a college preparatory program and this likely indicates a lower level of intelligence, these variables are expected to be negatively related to reservation wages.

Prior work experience is another indicator of the level of human capital an individual has accumulated. To reflect this, a variable is used indicating the percentage of the prior calendar year a person was employed (PCTWRK78). This indicates the recent work experience of a survey respondent and is expected to positively affect reservation wages.

Also, the percentage of the calendar year a respondent was employed after the survey year is used to capture the effect of unobservable human capital (PCTWRK80). This variable is expected to be higher for individual's with higher levels of human capital that cannot be quantified in terms of education or experience. Higher levels of human capital lead to a stronger attachment to the labor force because of the investment aspect of human capital. The more human capital a person has obtained, the stronger their commitment to the labor force. Accordingly, this variable is expected to affect reservation wages positively.

Clearly, raw intelligence affects human capital as much as, if not more than, education and experience. The only measure of intelligence available within the

NLSY is the results of a quiz concerning the nature of work in several different occupations. This quiz is referred to as the "Knowledge of and Experiences with the World of Work" test.

Respondents were asked what type job duties best fit nine different job titles. The variable constructed takes on integer values between 0 and 9 indicating the number of correct responses an individual had on the quiz (KNOWWORK). This is a rough measure of a person's intelligence, but it is expected to capture part of the effect of intelligence on reservation wages. The effect of this variable on individual's reservation wages is expected to be positive.

Numerous indicators of a person's attitudes exist. In the NLSY, respondents were asked a series of questions which served to provide a value for each respondent on what is referred to as the Rotter scale (ROTTER). The Rotter scale is a measure of an individual's beliefs that life's outcomes are within their control. The higher the score on the Rotter scale, the more a respondent believes they can affect their future.

Respondents who feel they have more control of their lives are likely to act in ways that are positively rewarded in the labor market. This indicates a higher mean of the available wage distribution. Accordingly, a higher Rotter score would be expected to produce higher reservation wages for an individual.

Labor market factors that affect the mean of the available wage distribution are the next logical points of discussion. The labor market factors used in this study include the unemployment rate (URATE), poverty rate (POORCNTY), and income level (PCAPINC). Higher levels of the local unemployment rate and poverty rate

indicate less attractive employment opportunities. In terms of the wage distribution, this is associated with lower expected wages and lower means of the available wage distributions. So, higher levels of unemployment and poverty, on a local level, are expected to result in lower reservation wages.

Per capita income is the specific measure used to indicate the local level of income. The expectation is that higher levels of income indicate a high wage market. A high wage market results in a higher mean wage of the available wage distribution. Thus, higher per capita income can be expected to lead to higher reservation wages.

Several different variables provide information about an individual's net cost of search. The first such variables are a respondent's labor force status. Dummy variables will indicate whether a respondent is employed (EMPLOYED) or out of the labor force (OLF). Unemployed is the category of labor force status that is omitted.

If a person is employed the wages that individual earns can also be expected to affect reservation wages. To capture this affect, the natural logarithm of the individual's current hourly wage (LOGWAGES) is included as an independent variable. The relationship of these variables to a worker's reservation wage is expected to be positive as explained in Chapter 2.

Also, whether or not a respondent is enrolled in regular school will affect the net search costs. A dummy variable for this category will be used in estimating reservation wages (NOTENRL). The dummy is constructed such that it takes the value 1 if a respondent is not enrolled in regular school at the time of the survey, 0

otherwise. Regular school is defined as elementary school, high school, college, or graduate school.

It should be noted, however, that being enrolled in school does not preclude a respondent from participation in the labor force. If an individual is not enrolled in regular school, this is expected to indicate a lower net search cost as explained in Chapter 2. This, in turn, indicates a lower reservation wage.

The last variables used to measure the opportunity cost of employment on reservation wages are indicators of receipt of welfare payments or willingness to accept such payments. A dummy variable is constructed which takes the value 1 if anyone in the respondent's household received welfare or public assistance in the year prior to the survey, 0 otherwise (WELFARE). To show the affect of willingness to accept welfare payments a second dummy variable is devised. This variable takes the value 1 if a respondent indicates in the survey that they "probably would accept welfare" if they were unable to support their family (ACPTWELF). If the respondent indicates they "probably would not accept welfare" even if they were unable to support their family, the dummy variable takes a value of 0. Both of these variables are expected to positively affect reservation wages.

Since the focus of this study is gender segregation, clearly the first personal characteristic variable should be an individual's gender. However, instead of using a dummy variable indicating a person's gender in estimation, a more general approach is utilized. For every job category, two log likelihood functions are estimated, one for female respondents and the other for male respondents.

A preliminary round of estimates of the log reservation wage equation, using a dummy variable for gender, showed that the gender dummy was significant in every job category. To allow for differences in the interaction of gender and other variables, separate equations are estimated for males and females. This is the reason Table 1 presents the means for the two sub-groups.

Other personal characteristics of the survey respondents that are used in the estimation of reservation wages include race, marital status, geographic region, parental background, and type of area in which the individual lives. There are no a priori expectations concerning the sign of the relationship between these variables and reservation wages since they represent personal characteristics. These variables are merely intended to capture any systematic differences in reservation wages across demographic groups.

Race, marital status, and region are indicated by dummy variables. The race categories used are Black (BLACK) and Hispanic (HISPANIC) with non-Black and non-Hispanic being the omitted category. For marital status, the dummy variable takes a value 1 if the individual is married, 0 otherwise (MSTAT). The categories of geographic region are North-central (NCENTRAL), South (SOUTH), and West (WEST). Northeast is the omitted category.

A variety of information is available concerning a respondent's parental background. The measure used in this study is to consider the education level of a respondent's parents. Specifically, the variable constructed reflects whether or not a person's parents graduated from high school. Two dummy variables are constructed

taking the value 1 if a respondent's mother/father graduated from high school, 0 otherwise (MOMSEDHS)/(DADSEDHS).

To indicate the type of area in which a respondent resides, a dummy variable approach could be used to indicate urban/suburban/rural areas. However, there is another approach that provides more information. A variable can be constructed measuring the degree of "urbanness" of the area in which a respondent lives. Specifically, this variable could be found as the percentage of the population in the respondent's area living in urban areas (URBAN). This may provide a better indicator of the type of area in which a respondent resides.

Having discussed the nature of the NLSY and the specific variables to be used in the estimation of reservation wages, a brief presentation of the steps taken to prepare the data for estimation is in order. Beginning with the full sample of 12,686 respondents, the first step in data processing is to remove any observation for which there are missing values for key variables. The variables for which a missing value results in an observation being dropped are the following: any of the questions concerning a hypothetical job offer (the dependent variable), highest grade completed (GRADE), percentage of weeks worked in 1978 (PCTWRK78), percentage of weeks worked in 1980 (PCTWRK80), any of the questions on the "Knowledge of World of Work" quiz (KNOWWORK), marital status (MSTAT), percent of local population in urban areas (URBAN), educational expectation (EDUCGOAL), labor force status (EMPLOYED)/(OLF), willingness to accept welfare (ACPTWELF), geographic

region (NCENTRAL)/(SOUTH)/(WEST), and Standard Metropolitan Statistical Area (SMSA).

Additionally, survey respondents who were on active duty in the military were dropped from the sample. This deletion is necessary since their responses to the questions concerning acceptance of a hypothetical job offer are affected by their current military status and would greatly differ from non-military survey respondents. Also, due to restrictions in the Current Population Survey concerning SMSA, as indicated in Chapter 4, respondents not residing in one of the forty-four SMSAs for which CPS data is available are deleted from the sample. Also, respondents under the age of 16 at the time of the survey were deleted from the sample. This step is necessary since only individuals 16 years of age and older were asked about their labor force status.

Several variables, for which there are a significant number of missing values, are handled differently than those above. For these variables a dummy variable is specified that takes the value 1 if the variable of interest has a missing value, 0 otherwise. These dummies are referred to as NA (information not available) variables.

The variables for which an NA dummy is constructed are the following: mother a high school graduate (MOMSEDNA), father a high school graduate (DADSEDNA), current wage (WAGENA), past acceptance of welfare (WELFNA), type of high school program (HSTYPENA), and Rotter scale (ROTTERNA). The reason for constructing these NA variables is to retain a sufficient number of

observations from which a distribution of workers is to be generated. Without this approach, as much as 35% of the usable observations could be lost because of missing values.

These deletions from the full NLSY sample and construction of NA variables leave 2976 observations which are split into the two samples, male and female, with 1441 and 1535 respondents respectively. These samples are hereafter referred to as the male and female samples. As previously stated, Table 3 shows the characteristics of these samples.

By examining Table 3 it can be seen that the average female was slightly older, had completed more education, and was more likely to have graduated from high school than the average male. Also, the average female worked less weeks in both 1978 and 1980 and was more likely to be married. On average, the females were more likely to be enrolled in regular school or out of the labor force and less likely to be employed than their male counterparts. Finally, the average wage appears to be lower for females than males on average, but this result could be caused by the lower level of employment for females in this sample.

The final step taken in preparing the NLSY data for the estimation process is to normalize the variables. That is, to aid in convergence of the maximum likelihood estimation, variables need to be of roughly the same magnitude. To accomplish this task, several variables are divided by some factor to standardize the data. The specific variables that are normalized in this fashion include: AGE, GRADE,

PCTWRK78, PCTWRK80, ROTTER, KNOWWORK, EDUCGOAL, URATE, URBAN, POORCNTY, and PCAPINC.

AGE is divided by 22 since the oldest survey respondent is 22 years of age. Both GRADE and EDUCGOAL are divided by 18 for similar reasons. ROTTER is divided by 4 and KNOWWORK is divided by 9 since these are the maximum scores attainable on each of these scales.

URATE, which is expressed with one implied decimal place, is divided by 132 which is the highest rate reported in the male and female samples. URBAN and POORCNTY are both divided by 1,000 since these variables are expressed with one implied decimal place. PCAPINC is divided by 10,000 since that variable takes on values between $2,220 and $8,493.

All of these transformations normalize these variables to be weakly between 0 and 1. With this normalization, all the variables used to estimate reservation wages are weakly between 0 and 1 except for the log of the current wage. When these steps are complete, the data is in the final form for estimation of the individual's reservation wages.

Having thoroughly covered the NLSY and the data used to estimate reservation wages, the discussion can now turn to the data for estimating wage distributions. As mentioned previously, the Current Population Survey, March 1979 will be used for this purpose. Before presenting the data used to estimate the relevant wage distributions, it seems prudent to discuss the method and nature of this survey.

The CPS is a monthly survey conducted by the Bureau of the Census for the Bureau of Labor Statistics. The primary purpose of the CPS is to provide policy-makers with information concerning the levels of employment and unemployment. This information is broken down into several different categories such as employment by industry, employment by sector (farm/non-farm), full-time/part-time employees, and total unemployment (U.S. Bureau of the Census, 1980).

The interviews for the CPS are conducted monthly and in 1979 approximately 55,000 households with over 100,000 individuals were surveyed, representing the non-institutionalized population of the United States. The format for the survey is a revolving 4-4-4 plan. This means that a household is not interviewed for four months, then interviewed the following four months. During the next four months, the household is again not interviewed.

The cycle is then repeated for each household. This format allows for not only month by month comparisons of the survey results, but also a year by year comparison. This is because from one month to the next, 75% of the households are the same and from one month to the same month in the next year, 50% of the households are the same. Each household in the CPS is surveyed only eight times, or over a two year period.

The survey in March of every year contains a supplemental questionnaire. In addition to the questions concerning labor force status individuals are also asked to provide information concerning their personal characteristics as well as their work history from the previous year. This supplement, known as the Annual Demographic

File, provides a yearly estimate of different characteristics of the entire population of the United States. A few of these characteristics include age, racial characteristics, gender composition, marital status, health, family structure, and education levels. Most importantly, the information provided by the Annual Demographic File represents not just the labor force but the entire non-institutionalized population of the United States (U.S. Bureau of the Census, 1980).

In addition to the standard demographic questions in the March supplement, numerous questions are included about income and work in the prior calendar year. A few of these questions can be used to estimate hourly wages for individuals. The specific survey information that can be used for this purpose is the individual's income earned from work in the past year, weeks worked in the past year, and average hours worked per week in the past year.

The hourly wage can then be found as the earnings from work in the prior year divided by the product of weeks worked in the past year and average hours worked per week. That is,

$$\text{Hourly Wage} = \frac{\text{Income earned from work in 1978}}{(\text{Weeks worked in 1978}) * (\text{Average Hours per Week in 1978})} \quad (1)$$

Observations for which any of these three variables had a missing value were dropped from the sample.

Also, only respondents who resided in a specific SMSA and were employed in one of the four specific job categories considered, were retained in the sample. Additionally, any estimated wage that was less than $2.25 per hour or greater than

$15.00 per hour was dropped from the sample since the minimum wage in 1978 was $2.25 and it seems unlikely that a person working in one of these four occupations would have earned more than $15 per hour in 1978. Finally, only respondents between the ages of 16 and 22 years of age, inclusive, were retained in the sample. This last exclusion is made in order to have a distribution of wages as comparable to the sample from the NLSY as possible.

When the hourly wage for the individuals in the CPS is constructed, the next step is to determine their occupation. This is found by using the information from the Annual Demographic File concerning their occupation in the prior calendar year. The occupation in which the individual was employed the longest in the prior year is assumed to be the occupation in which the respondent earned all of their income from work. This assumption is necessary since the survey does not separate the respondent's income by each job held in the prior year. This constructed hourly wage can then be used to estimate the necessary 176 wage distributions by the method described in Chapter 4.

There are several weaknesses with this approach to calculating the hourly wage. The first concern is caused by the nature of the data. The specific questions which provide the information used to calculate hourly wages were the following: "(Last calendar year, 1978,) How much did ... receive (in wages or salary) before any deductions?", "In 1978 how many weeks did ... work either full time or part time not counting work around the house? Include paid vacation and paid sick leave.", and "In

the weeks that ... worked, how many hours did ... usually work per week?". Since this information is self-reported, the reliability of the data may be questionable.

Of the responses used here, the most likely variable to be misrepresented by a respondent is income earned from work. The following passage from the Technical Documentation for the 1979 March CPS addresses this concern:

> "Moreover, readers should be aware that for many different reasons there is a tendency in household surveys for respondents to under report their income. From an analysis of independently derived income estimates, it has been determined that wages and salaries tend to be much better reported than such income types as public assistance, Social Security, and net income from interest, dividends, rents, etc." (U.S. Bureau of the Census, 1980)

This indicates that while a respondent's reported earned income from work may not be completely correct, it is less likely to be flawed than other types of income.

Another weakness of this construct for hourly wage is relying on the 'average' hours worked per week in the prior year. Since this average is multiplied by the number of weeks worked in 1978, a relatively small misrepresentation could greatly skew the hourly wage estimate. An additional problem with using the CPS data as described above is determination of an individual's occupation. It is assumed that every respondent earned all of their income from work in the occupation held for the longest in 1978. If a respondent held multiple jobs, or changed jobs, during the year this may bias the estimates of the hourly wage for that respondent and in turn, the occupation category. In spite of these potential flaws in the construction of the hourly wage, it will be used since no close substitute is available for estimating the necessary wage distributions.

This chapter has discussed at length the data used for estimation of reservation wages and wage distributions, the NLSY and CPS. Also, the variables used for estimation are discussed as well as the transformations of those variables, if any. All that remains for this study is to estimate both the reservation wages and wage distributions, match workers with occupations, and then examine the results of the segregation that would occur in the synthetic labor market developed in Chapter 3. This is the task undertaken in the next chapter, Chapter 6. The final chapter, Chapter 7, will present conclusions of the study and also discuss some potential improvements on this work for future studies.

TABLE 3
Sample Characteristics

Variable (1)	Description (2)	Means - Males (3)	Means - Females (4)
	Personal Characteristics		
AGE	Age, in years, at time of survey.	18.43	18.53
GRADE	Highest grade of regular school completed.	10.90	11.11
HSGRAD	Dummy variable taking the value 1 if respondent is a high school graduate.	0.3928	0.4599
HSVOCCOM	Type of high school curriculum. If	0.1520	0.1798
HSGENPRG	primarily vocational or commercial,	0.4650	0.4730
HSTYPENA	HSVOCCOM takes value 1. If primarily general program HSGENPRG takes value 1. If high school type is missing HSTYPENA takes value 1 and other variables take value 0. (Residual category is college prep.)	0.0673	0.0586
PCTWRK78	Percentage of weeks employed in 1978.	49.49	42.02
PCTWRK80	Percentage of weeks employed in 1980.	62.44	55.61
MSTAT	Dummy variable identifying if respondent is married.	0.0493	0.1303
BLACK	Dummy variable identifying respondent's	0.3248	0.3166
HISPANIC	racial-ethnic cohort (non-Black, non-Hispanic is the omitted category).	0.1957	0.1928

TABLE 3 (continued)
Sample Characteristics

Variable (1)	Description (2)	Means - Males (3)	Means - Females (4)
	Personal Characteristics		
MOMSEDHS MOMSEDNA	Highest grade completed by respondent's mother. If mother completed twelve or more years of regular school, then MOMSEDHS takes the value 1. If information not available, MOMSEDNA takes the value 1 and MOMSEDHS takes value 0. (Mother not a high school graduate is the omitted category)	0.5684 0.0604	0.5505 0.0599
DADSEDHS DADSEDNA	Highest grade completed by respondent's father. If father completed twelve or more years of regular school, then DADSEDHS takes the value 1. If information not available, DADSEDNA takes the value 1 and DADSEDHS takes value 0. (Father not a high school graduate is the omitted category)	0.5121 0.1506	0.5075 0.1668
WELFARE WELFNA	Acceptance of welfare in prior 12 months. If anyone in respondent's family received welfare during the prior 12 months WELFARE takes value 1. If information not available, WELFARE takes value 0 and WELFNA takes value 1.	0.1534 0.0285	0.1570 0.0189
ACPTWELF	Dummy variable equal to 1 if respondent indicates they "probably would accept welfare" if unable to support their family.	0.7037	0.6756

TABLE 3 (continued)
Sample Characteristics

Variable (1)	Description (2)	Means - Males (3)	Means - Females (4)
	Personal Characteristics		
ROTTER ROTTERNA	Number of correct responses on survey's "Rotter-scale" quiz indicating belief that life's outcomes lie within respondent's control (values range from 0 to 4). If information not available, ROTTER takes value 0 and ROTTERNA takes value 1.	2.677 0.0118	2.553 0.0052
KNOWWORK	Number of correct responses on survey's "knowledge of work" quiz (values range from 0 to 9).	5.981	5.855
EDUCGOAL	Highest grade respondent expects to complete.	13.85	13.94
NOTENRL	Dummy variable equal to 1 if respondent is not enrolled in school at time of survey.	0.4053	0.3557
EMPLOYED OLF	Dummy variable indicating labor force status of respondent during survey week (unemployed is the omitted category).	0.5635 0.2644	0.5010 0.3446
LOGWAGES	Log of the hourly wage, or 0 if hourly wage is missing or 0 if hourly wage is less than \$1 or greater than \$100[a].	0.6517	0.5045
WAGENA	Dummy variable equal to 1 if hourly wage is missing and person is employed.	0.0569	0.0469

TABLE 3 (continued)
Sample Characteristics

Variable (1)	Description (2)	Means - Males (3)	Means - Females (4)
	Labor Market Characteristics		
URATE	Continuous unemployment rate for labor market of current residence.	5.980	5.974
URBAN	Percent of population in county of current residence that lives in urban area.	92.31	92.57
POORCNTY	Percent of families in county of residence with incomes below the poverty line.	8.600	8.660
PCAPINC	Per capita income in county of current residence ($).	5083	5081
NCENTRAL	Dummy variable identifying geographical region of residence (Northeast is the residual category).	0.2866	0.2691
SOUTH		0.2241	0.2391
WEST		0.2047	0.1883
Observations		1441	1535

Notes: [A] LOGWAGES is set to zero if the reported hourly wage is less $1 per hour or greater than $100 per hour since observations beyond these bounds are not credible. The minimum wage in 1979 was $2.25 per hour. Since no worker could be legally earning less than this amount, a reported hourly wage below $1 per hour lacks any believability. Also, since the respondents in the male and female sub-samples are 22 years old or younger, but at least 16 years old, it seems equally unlikely that any of these respondents would be earning an hourly wage in excess of $100 per hour.

Chapter 6

EMPIRICAL RESULTS

Having discussed the model used for matching workers by tastes, the methods for this matching, and the data used, a presentation of the empirical results of this work is now in order. There are several sets of empirical results to be discussed. First, the results of the estimation of reservation wages need to be presented. Second, results from the estimation of wage distributions should be examined. Finally, the amount of segregation predicted by the synthetic labor market model will be discussed. The next, and final, chapter will present conclusions and extensions of this work.

The first step in the process by which workers are matched with occupations, based on tastes for the non-wage attributes of those jobs, is to estimate reservation wages. This estimation process utilizes several programs written to utilize the computer package GAUSS©. Different programs are used to estimate reservations wages for the male and female samples individually. Additionally, separate programs were written for each job category within the two samples. These programs can be found in Appendices A through H.

First, equation 3 from Chapter 4 is estimated to allow for an examination of the relationship between the independent variables and reservation wages. These results are presented in Tables 4 and 5 for the male and female estimations, respectively. The discussion of these results will focus on the variables that result in a significant parameter estimate in at least two of the four equations estimated.

The reason for examining parameter estimates from the simpler formulation of reservation wages found in equation 3 in Chapter 4 instead of the more complete version represented by equation 10 in Chapter 4 is that the coefficients of the latter equation are difficult to interpret. This is due to the inclusion of the dummy variables representing the identified category in which the reservation wage is located for the non-j job categories. These variables capture much of the effects expected to be shown by other explanatory variables. Accordingly, very few of the parameter estimates of equation 10 are significantly different from zero. The parameter estimates of equation 10 from Chapter 4, the full model of reservation wages, using all available information, are presented in Appendices I and J, for the male and female samples respectively.

Table 4 shows the parameter estimates for the reservation wage equations of males in the four different job categories. Variables reflecting the mean of the available wage distribution which produce significant estimates in at least two equations are AGE, GRADE, PCTWRK78, KNOWWORK, ROTTER, POORCNTY, PCTWRK80, and LOGWAGE. The first four of these represent standard human capital variables and yield positive estimates as theory predicts.

ROTTER identifies individuals who hold certain beliefs that are likely to be rewarded in the labor market with higher wages. The parameter estimate associated with this variable is positive, as expected. POORCNTY identifies the proportion of the respondent's county population that had incomes below the poverty level. This likely indicates a lower average wage in the respondent's area. Contrary to

expectations, the parameter estimate associated with POORCNTY is positive. PCTWRK80 is expected to capture the effects of unobserved human capital and its parameter estimate is positive as expected. LOGWAGE indicates the log of the current wage and produces a positive estimate as expected.

Variables reflecting the net cost of search which are significant in at least two of the four equations predicting reservation wages for males are the dummy variables OLF and NOTENRL. These variables serve to indicate whether a respondent is out of the labor force or not enrolled in regular school. OLF produces a positive estimate as expected. Also, the parameter estimate for NOTENRL is negative which is what theory predicts.

Only two personal characteristics produce estimates that are significant in half or more of the job categories. These are BLACK and SOUTH. Both of these variables produce negative estimates indicating a lower reservation wage, on average, for blacks or respondents living in the south. The only other variable that has significant parameter estimates in a majority of the job categories for males is ROTTERNA, which is positive. This dummy variable indicates whether a respondent has a missing value for one of the Rotter Scale questions. This result is unexpected, but it must be kept in mind that only 17 respondents had a value of 1 for this variable.

A correlation analysis provides some additional information about this peculiar result. ROTTERNA is positively correlated with the variables AGE, MSTAT, HISPANIC, and HSTYPENA with at least 90% confidence. Conversely,

ROTTERNA is negatively correlated to GRADE, DADSEDHS, KNOWWORK, EDUCGOAL, and NOTENRL with 90% or greater confidence. A picture emerges where these 17 respondents are older, have less education, are more likely married, are more likely to be Hispanic, have fathers that tend not to be high school graduates, scored poorly on the Knowledge of the World of Work Quiz, have lower educational goals, and tend to not know what type of high school program they studied. Based on this information a reasonable explanation of this result is the strong correlation between ROTTERNA and the variables AGE and NOTENRL.

The parameter estimates for the reservation wage equations in the four occupations for females are represented in Table 5. Variables reflecting the effect of the mean of the wage distribution that produce significant parameter estimates in at least 2 reservation wage equations include AGE, GRADE, HSGRAD, PCTWRK78, PCTWRK80, and PCAPINC. The sign of each of these estimates is positive, as expected.

The only variables representing net search cost that are consistently significant in two or more of the job category equations for the female sample are ACPTWELF, NOTENRL and OLF. ACPTWELF is a dummy variable representing a respondent's willingness to accept welfare if unable to support their family. The estimates for all three of these parameters are positive. These results are as expected for ACPTWELF and OLF. NOTENRL, however, is expected to negatively affect reservation wages, as was the case with the male sample. Clearly, not being enrolled in school affects reservation wages differently for males and females. One possible explanation for

this difference in sign between the male and female samples is that NOTENRL may be a proxy for some other variable that affects reservation wages differently for males and females.

Three personal characteristic variables result in parameter estimates that are significant in at least half of the equations for females. These are MSTAT, BLACK, and HISPANIC. As was the case for the male estimates, this parameter estimate for BLACK is negative. MSTAT and HISPANIC both produce positive estimates.

Considering both sets of estimations combined shows the following variables to be consistently significant in at least 4 of the 8 equations: AGE, GRADE, PCTWRK78, PCTWRK80, OLF, and BLACK. All of these variables produce signs that are consistent across gender samples. The variable BLACK merely captures differences in reservation wages between black respondents and non-black, non-Hispanic respondents (the omitted category). The negative sign for the parameter estimate associated with the variable BLACK indicates that black respondents have, on average, lower reservation wages. AGE, GRADE, and PCTWRK78 are human capital variables and all have positive predicted signs as theory predicts. OLF, the dummy variable indicating that a respondent is out of the labor force, also produces a positive parameter estimate, as expected.

The only variables for which the estimates are insignificant in all 8 equations are MOMSEDNA, DADSEDNA, WELFARE, and WELFNA. The only one of these that is curious is WELFARE. This variable is expected to positively affect the reservation wage since acceptance of welfare would lower the net cost of searching

for work. A possible explanation for this is that WELFARE may actually be capturing multiple effects. Persons with lower labor market potential are more likely to be welfare recipients. Thus, the variable WELFARE may actually be measuring both search costs and mean of the available wage distribution. Not coincidentally, these two factors work counter to each other, possibly explaining the lack of significance for this parameter estimate.

There are several variables that produce at least one significant estimate that contradicts theory. These variables are NOTENRL, URATE, URBAN, and POORCNTY. NOTENRL and URATE have opposite signs in the male and female equations indicating that these variables affect reservation wages differently by gender. For both of these variables, the estimate in the female equation contradicts the predictions of theory. URBAN and POORCNTY are consistent in opposing the predictions of theory. Clearly, these variables are capturing effects other than those expected.

The discussion can now turn to the estimation of the 176 wage distributions. Estimation of equation (12) in Chapter 4 is accomplished by OLS and the results are presented in Table 6. A positive estimate indicates a higher mean wage in that SMSA than in the base SMSA, which is Washington, D.C. -Maryland -Virginia. A negative estimate indicates a lower mean wage in that SMSA relative to the base SMSA. Job category dummy variables were also used in this regression of wages. These parameter estimates are not included in Table 6 since only the effect of a specific SMSA on the distribution of wages is needed.

For SMSA 6040, Patterson-Clifton-Passaic, New Jersey, no estimate could be obtained. This is due to a lack of observations in this SMSA in any of the four job categories for individuals between the ages of 16 and 22. There is an alternative to not estimating the parameter for this SMSA.

The alternative is to include the dummy variables for SMSA 8840, Washington D.C., and run the regression without an intercept. While this option is econometrically sound, it does not satisfy the needs of the current model. The reason this regression is needed is to obtain values to shift the means of the four occupation specific wage distributions. A regression without an intercept and no excluded categories would merely indicate the average wage, across the four job categories, within each SMSA.

This information could be subtracted from the overall average wage in all SMSAs and occupations to derive the necessary 'mean-shifters'. However, such a procedure would still provide no information regarding how the mean wage in the four job categories differs from the overall average in SMSA 6040, Patterson-Clifton-Passaic, New Jersey. Thus, this alternative does not solve the problem. The only remaining alternative then is to assume that the average wages in SMSA 6040 exactly match the average wages in the omitted SMSA, Washington, D.C. –Maryland – Virginia.

The parameter estimates in Table 6 are significant for only 4 of the 42 SMSAs. It might be expected that the estimates would more likely be significant for larger SMSAs. To consider this effect, the size rank of each SMSA is reported.

These rankings range from 1 to 44. A smaller size rank indicates a larger SMSA in terms of observations in the sample.

The size ranks of the four SMSAs with significant parameter estimates are 8, 13, 31, and 41. This information seems to indicate that SMSA size is, most likely, not a major factor affecting the significance of the estimate. A point of interest here is the fact that all of the significant estimates are positive. In fact, only 17 of the 42 estimates are negative.

Each of the estimates in Table 6 are used to shift the wage distributions shown in Chapter 5 by the amount of the estimate in Table 6 for the specified SMSAs, regardless of significance. A positive estimate results in a rightward shift of the four distributions of wages for that particular SMSA. A negative estimate results in a corresponding leftward shift of the wage distributions for each of the 4 job categories in that SMSA.

For example, in SMSA 5600, representing the city of New York, each of the log-wage distributions shown in Chapter 5 are shifted to the right by 0.0011. Correspondingly, in SMSA 5640, Newark, NJ, the log-wage distributions are shifted to the left by 0.2546. This process generates the 176 wage distributions by occupation-SMSA pairs.

Now that an individual's reservation wages in several job categories and the corresponding wage distributions for those job categories have been found, workers can be matched with one of the four specific occupations. This matching process is carried out in five different methods as discussed in Chapter 3. The results of these

different matching methods are presented in Tables 7 through 10. The discussion that follows will briefly review each method by which workers are sorted and then examine the results of that particular matching of workers with occupations.

The first method for matching workers with job categories assumes that the probability of receiving a job offer is constant across occupations. The probability of accepting a job offer is calculated using the reservation wage and the distribution of wages for each job by SMSA. When this value is calculated for each worker in the four job categories, workers are matched with the occupation where their probability of accepting an offer is the greatest. These results are shown in Table 7. The overall proportion of females in the full sample, the summation of the male and female samples, is 0.5158. The predicted proportions of females in each of the four job categories are presented in column 3 of the table.

This method for sorting workers shows that 57.9% of those matched with the occupation Dishes and 47.5% of the workers matched with the job Burgers are female. These proportions are significantly different than the proportion of females in the full sample. Of the workers matched with the occupations Cleaning and Supermarket, 49.6% and 53.4% respectively, are female. These proportions are not significantly different from the overall proportion of females in the full sample.

Based on these results, it seems clear that differences in preferences for work characteristics between genders do affect the distribution of workers across jobs. Additional evidence supporting this finding is the result of a Chi-squared test for independence of occupational outcome and gender in method 1. The Chi-squared

statistic for this test is 17.623 which indicates that gender and occupation are dependent with 99.5% confidence.

The second method for matching workers with one of the four specific occupations is very similar to the first. The only difference is how workers are placed into occupations. The first match placed workers entirely into the occupation where the probability of accepting a job offer was the greatest. The second method will place workers into job categories according to the probability of accepting a job offer relative to the sum of this probability for the individual for all jobs in the model. That is, workers will be p-sorted by their probability of accepting job offers, found by using reservation wages and the wage distributions, assuming the probability of receiving an offer is constant across all occupations for each worker. These results are presented in Table 8.

The level of segregation predicted by this second method of matching workers is far less than that predicted by the first method. In fact, the distribution of workers in all four job categories almost exactly matches the overall distribution in the sample. For example, with this matching method, 53.5% of the respondents matched with the job Dishes are female. The proportion of females in the job categories Cleaning, Supermarket, and Burgers are 50.1%, 51.2%, and 52.4% respectively.

The proportion of females in each job category differs from the overall proportion of females by less than two percentage points. This set of results indicate that taste differences between genders for the attributes of a job are not a significant cause of gender segregation. The Chi-squared statistic for the test of independence

between gender and occupation has a value of 2.180. The result of this test is that gender and occupation are not found to be dependent at any conventional level of significance.

The next two methods for matching workers with occupations use more information than the preceding methods. Specifically, the following methods include estimates of the probability of receiving a job offer in each of the four job categories. These probabilities are estimated based on the relative number of persons employed in each occupation and are shown in Table 11.

For the third method by which workers are sorted, the probability of a worker accepting a job offer is estimated, as before, using reservation wages and wage distributions. This probability is then multiplied by the probability of receiving an offer in the given job category resulting in the probability of a worker becoming employed in that occupation. Workers are then sorted into the occupation where they have the highest probability of being employed. The results of this match are shown in Table 9.

The probability of a worker receiving a job offer is relatively small for the job categories Dishes and Burgers. This results in a very small number of workers being matched with these two occupations. Specifically, 13 workers are matched with the job Dishes, 8 of which are female, and 53 workers are matched with the job Burgers, 29 of which are female. This extremely small number of observations prevents testing for a difference between the proportion of females in these occupations and

the overall proportion of females in the sample. Thus, no inferences about these specific occupations can be drawn.

The remaining two occupations, Cleaning and Supermarket, both have relatively large numbers of workers, 2620 and 290 respectively. The proportion of females in Cleaning is 51.1% and in Supermarket is 55.2%. Neither of these proportions of females differ significantly from the proportion in the full sample. Testing for independence of gender and occupation results in a Chi-squared statistic of 2.498 indicating that independence cannot be rejected at even the 90% level of confidence.

The fourth, and final, method of matching workers with occupations improves on the third method by p-sorting the workers. As before, the probability of accepting a job offer is found by using reservation wages and the appropriate wage distribution. The probability of receiving a job offer is estimated as described for the fourth method and these probabilities are listed in Table 11. The results of this final matching of workers to occupations are presented in Table 10.

As would be expected, since workers are p-sorted, the fourth match produces less segregation than the third. As was the case in the third method, the categories Dishes and Burgers have a relatively small number of workers, 189.0 and 294.4 respectively, because the probability of receiving an offer in these categories is relatively small. However, in this fourth match of workers, the number of observations in these job categories is sufficient to allow testing for difference of proportions.

The results of such tests show that in neither occupation does the proportion of females, 54.2% for Dishes and 52.7% for Burgers, differ significantly from the proportion in the full sample. The remaining occupations, Cleaning and Supermarket, again have a relatively large number of workers and the proportion of females in both categories, 51.4% and 50.8% respectively, closely mirrors the overall proportion. Finally, testing for independence of occupation and gender produces a Chi-squared value of 0.823 indicating, again, that independence cannot be rejected at any of the conventional levels of confidence.

An interesting exercise is to examine the predicted segregation in the occupations across each sorting method. The discussion in this area will focus on the outcomes of the different distributions, the results of which are combined in Table 12 As previously stated, any conclusions that can be drawn from these comparisons will be discussed in the next, and final, chapter.

The jobs category Dishes displays results that are unique within this work. In the first match dishes is a female dominated job category, as indicated by the predicted proportion of females of 57.9% which is greater than the overall proportion of females of 51.9%. This result is also significant. The final three matches also indicate that the occupational category dishes is female dominated, resulting in proportions of females in the job Dishes of 53.5%, 61.5%, and 54.2% for methods two through four, but these results are not statistically significant. What makes this occupation unique is that all four of these outcomes contradict the segregation

observed in the actual labor market. In reality, Dishes is a male dominated job while the model here predicts in all cases that it is female dominated.

The job category Cleaning produces less curious results than the other three occupations. All four methods for matching workers with occupations predict cleaning to be a male dominated job category without significance. The proportion of workers predicted in this job that are female are 49.6%, 50.1%, 51.1%, and 51.4% for the four different methods. Thus, cleaning is consistently a male dominated occupation in all four sorting methods, a result that corresponds well with observed segregation. However, none of the results for this job differ significantly from the overall sample.

Supermarket is a category that has somewhat curious outcomes. In the first sort 53.4% of the workers are female, indicating it is a female dominated occupation. This proportion is significantly different from the proportion of females in the full sample. In none of the remaining matches does supermarket produce a statistically significant outcome. However, it does change in female or male dominance. In the third match, supermarket is again a female dominated job category with 55.2% of the workers being female. In the second and fourth matches, however, Supermarket is male dominated with 51.2% and 50.8% of the workers being female. These results are especially interesting when contrasted to the actual segregation indicated by the fact that 87.1% of those working in this job category are females.

In the final job category, Burgers, peculiar results are again found. In the first match of workers to jobs, burgers is a male dominated occupation with a proportion

of females in the job of 47.5%. This result is statistically significant. In the second and fourth matches, burgers is a female dominated job category with 52.4% and 52.7% females respectively. However, neither of these results is significant. In the third method, burgers is predicted to be perfectly integrated. That is, the proportion of females in this case exactly matches the proportion in the overall sample. The only result that is significant is that where burgers is male dominated. This may be of significance since, in reality, burgers is a male dominated occupation since on 38.1% of those employed in this job are females.

Examination of these results indicates that taste differences between workers may, or may not, be a significant determinant of gender segregation. The next logical question is how well the results predicted here can explain the observed occupational gender segregation in all four job categories. That is, what portion of the observed total level of segregation is explained by this synthetic labor market where the occupation of a worker is determined primarily by that worker's tastes for different jobs. To estimate the amount of total segregation explained by this model a measure is needed that quantifies aggregate segregation across the different job categories.

To measure the level of overall segregation in all four job categories the index of dissimilarity, S_j, will be used where:

$$S_j = \frac{1}{2}\sum_{i=1}^{4}|M_{ij} - F_{ij}| \qquad (2)$$

where M_{ij} is the proportion of total males in occupation i predicted by matching method j and F_{ij} is the corresponding value for females. The index of dissimilarity is

weakly bounded by zero and one. A value of zero indicates that the work-force is perfectly integrated while a value of one indicates that it is perfectly segregated. An interpretation of this measure is that it shows the proportion of males, or females, that would have to change jobs for the work-force to be perfectly integrated (Duncan and Duncan, 1955). The index of dissimilarity for the actual distribution of workers in the four occupations considered here is 0.4068 indicating that just over 40% of males, or females, would have to change jobs to reach perfect integration in these four jobs.

Column 2 of Table 13 shows the index of dissimilarity for each of the five different methods of matching workers with jobs. These values range from 0.0108 to 0.0684. To measure the amount of actual segregation that can be explained by taste differences a ratio of these values to the actual index of dissimilarity for the four job categories can be used. These values are reported in column 3 of Table 13 and represent the portion of actual total segregation that is predicted in each of the four different matches of workers.

Table 13 indicates that between 2.67% and 16.81% of the observed segregation in these four job categories can be attributed to individuals self selecting due to differences in preferences. This measure, however, should be used with a note of caution. In every method used to sort workers, at least one job category produces segregation that directly contradicts observed segregation. That is, a job category is predicted to be female (male) dominated when, in fact, it is male (female) dominated. Also, it should be remembered that in the last three matches independence between gender and occupation cannot be rejected.

This chapter has discussed several levels of results. First, the estimates from the reservation wage equation, equation 3 in Chapter 4, were presented. Next, wage distributions, for the four different job categories in 44 different geographic locations, were estimated. Finally, the reservation wages were utilized with the wage distributions to generate a distribution of workers. This last step was repeated five times under different assumptions concerning the amount of information affecting the match of workers to jobs. These generated distributions were then compared to actual distributions of workers. All that remains is an interpretation of these results and general conclusions. It is this task to which the final chapter is devoted.

TABLE 4
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
INTERCEPT	-0.1775	-0.1633	0.2147	-0.5475
	(0.515)	(0.493)	(0.621)	(1.511)
AGE	0.6255**	0.6291**	0.9442***	0.5789*
	(2.089)	(2.185)	(2.994)	(1.815)
GRADE	0.5776*	0.5837**	0.6538*	1.5205***
	(1.957)	(2.078)	(1.868)	(4.841)
HSGRAD	0.0297	0.0071	0.0127	-0.0927
	(0.498)	(0.124)	(0.11)	(1.461)
HSVOCCOM	0.0303	0.0269	0.0439	0.0435
	(0.626)	(0.568)	(0.846)	(0.848)
HSGENPRG	0.0211	-0.0017	0.077*	0.0617
	(0.548)	(0.052)	(1.869)	(1.54)
PCTWRK78	0.0955**	0.0642	0.0965**	0.1337**
	(1.998)	(1.417)	(2.025)	(2.571)
PCTWRK80	0.1519***	0.1108***	0.0663	0.0768
	(3.465)	(2.66)	(1.524)	(1.616)
MSTAT	0.0264	0.1027	0.0505	0.1201
	(0.372)	(1.246)	(0.614)	(1.447)
BLACK	-0.1737***	-0.1534***	-0.2236***	-0.2143***
	(4.570)	(4.017)	(5.856)	(5.229)

TABLE 4 (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
HISPANIC	0.0348	0.044	-0.0387	0.0602
	(0.784)	(0.891)	(0.876)	(1.243)
MOMSEDHS	-0.0133	0.0342	0.0284	0.0727*
	(0.368)	(1.014)	(0.761)	(1.876)
MOMSEDNA	-0.0420	0.06	-0.0422	0.104
	(0.648)	(0.849)	(0.647)	(1.453)
DADSEDHS	0.0765**	0.0335	-0.0222	0.0272
	(2.047)	(0.999)	(0.514)	(0.678)
DADSEDNA	-0.0114	-0.0308	-0.0214	-0.0176
	(0.249)	(0.603)	(0.375)	(0.358)
WELFARE	-0.0372	0.0272	-0.0427	-0.0254
	(0.870)	(0.569)	(0.918)	(0.549)
ACPTWELF	0.0152	0.0118	-0.0283	0.0159
	(0.491)	(0.439)	(0.869)	(0.466)
ROTTER	0.0909	0.0577	0.116**	0.1132*
	(1.612)	(1.079)	(2.015)	(1.852)
KNOWWORK	0.0881	0.2018***	0.1597**	0.2397***
	(1.230)	(2.831)	(2.233)	(3.071)
EDUCGOAL	0.1738	0.3073**	-0.1532	-0.2125
	(1.192)	(2.202)	(1.027)	(1.357)

TABLE 4 (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
NOTENRL	-0.0968*	-0.0494	-0.1033*	-0.1397***
	(1.892)	(1.05)	(1.721)	(2.595)
EMPLOYED	-0.0161	-0.0309	0.0304	-0.0549
	(0.189)	(0.424)	(0.406)	(0.637)
OLF	0.1038**	0.1437***	0.1686***	0.1578***
	(2.313)	(3.079)	(3.735)	(3.239)
LOGWAGES	0.1915***	0.1951***	0.1386***	0.226***
	(3.376)	(3.92)	(2.755)	(3.885)
WAGENA	0.1259	0.1135	0.125	0.1496
	(1.324)	(1.236)	(1.407)	(1.523)
WELFNA	0.1093	-0.0375	-0.0625	0.015
	(1.273)	(0.402)	(0.727)	(0.164)
HSTYPENA	-0.0188	0.0724	0.0782	0.0953
	(0.242)	(0.934)	(1.033)	(1.209)
ROTTERNA	0.2654*	0.1333	0.4447***	0.4136**
	(1.847)	(0.883)	(2.959)	(2.528)
URATE	-0.1624	0.0091	-0.3367*	-0.2344
	(0.909)	(0.052)	(1.704)	(1.168)
URBAN	0.0135	-0.0016	-0.0474	-0.1144
	(0.116)	(0.015)	(0.361)	(0.905)

TABLE 4 (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
POORCNTY	1.8699***	0.2753	0.1096	1.8095***
	(3.078)	(0.438)	(0.101)	(2.736)
PCAPINC	0.2386	0.1124	-0.2283	0.5208
	(0.758)	(0.364)	(0.584)	(1.529)
NCENTRAL	0.0636	0.0361	0.0193	0.0316
	(1.515)	(0.833)	(0.418)	(0.669)
SOUTH	-0.1227**	-0.0678	-0.1347*	-0.1261**
	(2.165)	(1.161)	(1.82)	(1.997)
WEST	0.0039	-0.0005	-0.0562	0.0217
	(0.086)	(0.011)	(1.146)	(0.401)
SCALE(σ_J)	0.2790***	0.2996***	0.2787***	0.3006***
	(28.465)	(31.181)	(28.703)	(27.498)
Pseudo R^2	0.1160	0.0868	0.1200	0.1308
Observations	1441	1441	1441	1441

Notes: Parameter estimates are significantly different from zero with the following levels of significance: *** 1% significance, ** 5% significance, * 10% significance

Pseudo R^2 is calculated as 1-[(log L_Ω)/(log L_ω)], where L_Ω is the maximized value of the unrestricted likelihood function and L_ω is the maximized value of the likelihood function with no explanatory variables (Maddala, 1983, p. 40).

TABLE 5
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
INTERCEPT	-0.4256 (1.175)	-0.4764 (1.313)	-0.6139** (2.344)	-1.5046*** (3.946)
AGE	0.2611 (0.785)	0.2939 (0.886)	0.8813*** (3.728)	1.5152*** (4.349)
GRADE	0.4776 (1.447)	0.6041* (1.823)	0.2964 (1.267)	0.6066* (1.723)
HSGRAD	0.2591*** (3.857)	0.0917 (1.375)	0.1901*** (3.633)	0.1073 (1.503)
HSVOCCOM	0.0273 (0.543)	0.0897* (1.768)	0.0438 (1.047)	0.0703 (1.359)
HSGENPRG	-0.0645 (1.593)	-0.0435 (1.074)	0.0155 (0.536)	-0.0009 (0.022)
PCTWRK78	0.1161** (2.136)	0.1778*** (3.252)	0.0615* (1.688)	0.2128*** (3.768)
PCTWRK80	0.1061** (2.241)	0.0564 (1.189)	0.0733** (2.198)	0.0761 (1.537)
MSTAT	0.1214** (2.184)	0.0449 (0.821)	0.0544 (1.286)	0.12** (2.123)
BLACK	-0.1544*** (3.617)	-0.0683 (1.572)	-0.0979*** (3.029)	-0.154*** (3.431)

TABLE 5 (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
HISPANIC	0.1098**	0.1632***	0.0262	0.0345
	(2.154)	(3.184)	(0.645)	(0.669)
MOMSEDHS	-0.0093	-0.0049	0.0053	0.0096
	(0.239)	(0.121)	(0.203)	(0.248)
MOMSEDNA	0.0291	0.0886	-0.0771	0.0473
	(0.405)	(1.22)	(1.244)	(0.624)
DADSEDHS	-0.0113	-0.0036	-0.0295	0.0313
	(0.27)	(0.095)	(1.123)	(0.761)
DADSEDNA	0.0212	0.0075	-0.0068	-0.0229
	(0.435)	(0.148)	(0.158)	(0.438)
WELFARE	-0.0468	-0.0653	-0.0594	-0.0549
	(0.99)	(1.387)	(1.424)	(1.101)
ACPTWELF	0.0761**	0.092***	0.0289	0.0856**
	(2.223)	(2.702)	(1.287)	(2.385)
ROTTER	0.0867	0.0849	0.0129	0.1698***
	(1.394)	(1.367)	(0.301)	(2.624)
KNOWWORK	0.2156***	0.0663	0.0584	0.1328
	(2.673)	(0.826)	(1.018)	(1.576)
EDUCGOAL	0.2395	0.2975*	0.0952	0.0734
	(1.442)	(1.789)	(0.803)	(0.422)

TABLE 5 (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
NOTENRL	0.1058*	0.04	0.0884*	0.0565
	(1.814)	(0.683)	(1.902)	(0.912)
EMPLOYED	0.1149	-0.0326	0.098*	0.0939
	(1.643)	(0.453)	(1.871)	(1.255)
OLF	0.0713	0.1078**	0.1341***	0.1806***
	(1.501)	(2.269)	(3.554)	(3.598)
LOGWAGES	0.0029	0.1047**	0.0257	0.035
	(0.06)	(2.166)	(0.711)	(0.684)
WAGENA	0.0724	0.2219**	0.0752	0.1246
	(0.756)	(2.294)	(1.006)	(1.261)
WELFNA	0.036	0.2124	0.1086	0.0991
	(0.298)	(1.619)	(1.048)	(0.78)
HSTYPENA	0.108	0.0531	0.1492**	0.1384
	(1.25)	(0.612)	(2.208)	(1.543)
ROTTERNA	0.5224**	0.3306	0.1493	0.3616
	(2.153)	(1.368)	(0.795)	(1.61)
URATE	0.1194	0.5596***	0.1877	0.2098
	(0.587)	(2.772)	(1.298)	(0.978)
URBAN	0.0972	0.0253	0.2092**	0.1956
	(0.774)	(0.196)	(2.164)	(1.459)

TABLE 5 (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation 3 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
POORCNTY	1.8511***	1.0676	0.478	0.4609
	(2.756)	(1.602)	(0.95)	(0.67)
PCAPINC	0.787**	0.7375**	0.1876	0.3352
	(2.335)	(2.141)	(0.791)	(0.959)
NCENTRAL	0.0047	0.019	0.0099	0.103**
	(0.097)	(0.407)	(0.276)	(2.042)
SOUTH	-0.0377	0.0455	-0.0161	0.1065
	(0.601)	(0.742)	(0.357)	(1.634)
WEST	-0.011	0.0589	-0.0103	0.1759***
	(0.195)	(1.071)	(0.241)	(3.019)
SCALE(σ_J)	0.3183***	0.3177***	0.26***	0.3386***
	(27.796)	(27.876)	(35.56)	(28.022)
Pseudo R^2	0.0821	0.0606	0.0667	0.0884
Observations	1535	1535	1535	1535

Notes: Parameter estimates are significantly different from zero with the following levels of significance: *** 1% significance, ** 5% significance, * 10% significance

Pseudo R^2 is calculated as 1-[(log L_Ω)/(log L_ω)], where L_Ω is the maximized value of the unrestricted likelihood function and L_ω is the maximized value of the likelihood function with no explanatory variables (Maddala, 1983, p. 40).

TABLE 6
Estimates of the Affect of SMSA on Wages, Equation 12 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses

SMSA (1)	Location (2)	Size Rank (3)	Estimate (4)
0080	Akron, OH	43	-0.1063 (0.407)
0160	Albany-Schenectady-Troy, NY	42	0.1294 (0.674)
0360	Anaheim-Santa Ana-Garden Grove, CA	17	-0.1451 (0.765)
0520	Atlanta, GA	14	-0.2348 (1.242)
0720	Baltimore, MD	13	0.3188* (2.085)
1000	Birmingham, AL	40	-0.1629 (0.861)
1120	Boston, MS	7	-0.0463 (0.352)
1280	Buffalo, NY	37	0.0037 (0.017)
1600	Chicago, IL	3	0.0552 (0.508)
1640	Cincinnati, OH	27	0.1112 (0.684)

TABLE 6 (continued)
Estimates of the Affect of SMSA on Wages, Equation 12 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

SMSA (1)	Location (2)	Size Rank (3)	Estimate (4)
1680	Cleveland, OH	18	-0.2165 (1.144)
1840	Columbus, OH	32	-0.1321 (0.945)
1920	Dallas, TX	16	-0.0332 (0.238)
2080	Denver, CO	26	0.0996 (0.686)
2160	Detroit, MI	5	0.0427 (0.404)
2800	Fort Worth, TX	38	0.3105 (0.883)
2960	Gary-Hammond-East Chicago, IN	44	0.0678 (0.316)
3120	Greensboro-Winston-Salem-High Point, NC	41	0.4518* (2.110)
3360	Houston, TX	9	-0.1147 (0.753)
3480	Indianapolis, IN	33	0.0027 (0.021)

TABLE 6 (continued)
Estimates of the Affect of SMSA on Wages, Equation 12 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

SMSA (1)	Location (2)	Size Rank (3)	Estimate (4)
3760	Kansas City, MO -KS	25	0.0988 (0.383)
4480	Los Angeles-Long Beach, CA	2	-0.0006 (0.006)
5000	Miami, FL	21	-0.1571 (0.737)
5080	Milwaukee, WI	28	-0.1754 (1.301)
5120	Minneapolis-St. Paul, MN	15	0.1318 (0.938)
5380	Nassau-Suffolk, NY	10	0.0588 (0.422)
5560	New Orleans, LA	31	0.4576* (1.798)
5600	New York, NY	1	0.0011 (0.010)
5640	Newark, NJ	19	-0.2546 (1.194)
5720	Norfolk-Portsmouth, VA	34	-0.0225 (0.106)

TABLE 6 (continued)
Estimates of the Affect of SMSA on Wages, Equation 12 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

SMSA (1)	Location (2)	Size Rank (3)	Estimate (4)
6040	Patterson-Clifton-Passaic, NJ	30	Not Estimated
6160	Philadelphia, PA	4	0.0347 (0.302)
6280	Pittsburgh, PA	12	0.0787 (0.455)
6440	Portland, OR -WA	35	0.5522 (1.572)
6840	Rochester, NY	39	-0.0054 (0.036)
6920	Sacramento, CA	36	0.3437 (1.603)
7040	St. Louis, MO -IL	11	0.0117 (0.080)
7280	San Bernardino-Riverside-Ontario, CA	24	-0.2100 (0.964)
7320	San Diego, CA	20	0.0410 (0.161)
7360	San Francisco-Oakland, CA	8	0.2492* (2.040)

TABLE 6 (continued)
Estimates of the Affect of SMSA on Wages, Equation 12 in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

SMSA (1)	Location (2)	Size Rank (3)	Estimate (4)
7400	San Jose, CA	29	0.0086 (0.045)
7600	Seattle-Everett, WA	23	0.1018 (0.721)
8280	Tampa-St. Petersburg, FL	22	-0.1656 (1.025)
8840	Washington, D.C. -MD -VA	6	Omitted Category
	Observations		351
	R^2		0.1405

Note: * Indicates the parameter estimate is significantly different from zero at the 10% significance level.

Table 7
Distribution of Workers for Method 1

Job Category (1)	Number of Predicted Workers in Occupation that are Female (2)	Proportion of Predicted Workers in Occupation that are Female (3)	Total Number of Predicted Workers in Occupation (4)
Dishes	366	0.5791***	632
Cleaning	488	0.4964	983
Supermarket	311	0.5344	582
Burgers	370	0.4750**	779
Total	1535	0.5158	2976

Notes: Predicted proportions of females in a job category are different than the proportion of females in the full sample at the following levels of significance:

* 90% confidence

** 95% confidence

*** 99% confidence.

Testing for independence of worker gender and occupation yields $\chi^2 = 17.623$. With 99.5% confidence, independence can be rejected.

Method 1 assumes $P(R_{ij})$ constant across occupations, estimates $P(A_{ij}|R_{ij})$ using reservation wage and wage distribution, then distributes workers with 0/1 sorting.

Table 8
Distribution of Workers for Method 2

Job Category (1)	Number of Predicted Workers in Occupation that are Female (2)	Proportion of Predicted Workers in Occupation that are Female (3)	Total Number of Predicted Workers in Occupation (4)
Dishes	385.230	0.5346	720.581
Cleaning	524.145	0.5006	1046.939
Supermarket	302.237	0.5115	590.903
Burgers	323.387	0.5236	617.576
Total	1535	0.5158	2976

Notes: Predicted proportions of females in a job category are different than the proportion of females in the full sample at the following levels of significance:

* 90% confidence
** 95% confidence
*** 99% confidence.

Testing for independence of worker gender and occupation yields $\chi^2 = 2.180$. Independence cannot be rejected at any acceptable level of significance.

Method 2 assumes $P(R_{ij})$ constant across occupations, estimates $P(A_{ij}|R_{ij})$ using reservation wage and wage distribution, then distributes workers with p-sorting.

Table 9
Distribution of Workers for Method 3

Job Category (1)	Number of Predicted Workers in Occupation that are Female (2)	Proportion of Predicted Workers in Occupation that are Female (3)	Total Number of Predicted Workers in Occupation (4)
Dishes	8	0.6154[a]	13
Cleaning	1338	0.5107	2620
Supermarket	160	0.5517	290
Burgers	29	0.5158[a]	53
Total	1535	0.5158	2976

Notes: Predicted proportions of females in a job category are different than the proportion of females in the full sample at the following levels of significance:

* 90% confidence
** 95% confidence
*** 99% confidence.

Testing for independence of worker gender and occupation yields $\chi^2 = 2.498$. Independence cannot be rejected at any acceptable level of significance.

[a] The proportions for these categories could not be tested due to insufficient observations.

Method 3 estimates $P(R_{ij})$ using the relative frequency approach, estimates $P(A_{ij}|R_{ij})$ using reservation wage and wage distribution, then distributes workers with 0/1 sorting.

Table 10
Distribution of Workers for Method 4

Job Category (1)	Number of Predicted Workers in Occupation that are Female (2)	Proportion of Predicted Workers in Occupation that are Female (3)	Total Number of Predicted Workers in Occupation (4)
Dishes	102.338	0.5415	188.997
Cleaning	895.953	0.5144	1741.706
Supermarket	381.653	0.5083	750.895
Burgers	155.056	0.5267	294.401
Total	1535	0.5158	2976

Notes: Predicted proportions of females in a job category are different than the proportion of females in the full sample at the following levels of significance:
* 90% confidence
** 95% confidence
*** 99% confidence.

Testing for independence of worker gender and occupation yields $\chi^2 = 0.823$. Independence cannot be rejected at any acceptable level of significance.

Method 4 estimates $P(R_{ij})$ using the relative frequency approach, estimates $P(A_{ij}|R_{ij})$ using reservation wage and wage distribution, then distributes workers with p-sorting.

Table 11
Actual Distribution of Workers

Job Category (1)	Proportion of Workers in Occupation that are Female[a] (2)	Total Number of Workers in Occupation (in Thousands) (3)	Probability of a worker receiving a job offer in the job category[d] (4)
Dishes	0.342	246,000	0.0565
Cleaning	0.436[b]	2,203,000	0.5062
Supermarket	0.871	1,434,000	0.3295
Burgers	0.381	469,000	0.1078
Total	0.5680[c]	4,352,000	1.000

Notes: [a] The proportion of workers in each occupation that are female was taken from the Current Population Survey for 1978. The categories from the CPS and the 3-digit occupation number that were matched to the occupations are as follows: DISHES - "dishwashers" (913); CLEANING - "chambermaids and maids, excluding private household" (901), "cleaners and charwomen" (902), "janitors and sextons" (903), and "maids and servants, private household" (984); SUPERMARKET" - "cashier" (310); BURGERS - "food counter and fountain workers" (914).

[b] The proportion of females for cleaning was calculated as a weighted mean of the percentage female in each of the CPS occupations where the number of persons employed in each occupation was used as the weight.

[c] The total proportion of females for the CPS data was calculated as a weighted mean of the percentage female in each job category where the number of persons employed in each occupation was used as the weight.

[d] The probability of a worker receiving a job offer in a job category is found as the proportion of workers in that category relative to the total number of workers in all job categories considered in this study.

Table 12
Comparison of Actual Segregation and Predicted Segregation

	Proportion of workers in occupation that are female.				
Job Category (1)	Actual (2)	Method 1 (3)	Method 2 (4)	Method 3 (5)	Method 4 (6)
Dishes	0.342	0.5791***	0.5346	0.6154	0.5415
Cleaning	0.436	0.4964	0.5006	0.5107	0.5144
Supermarket	0.871	0.5344	0.5115	0.5517	0.5083
Burgers	0.381	0.4750**	0.5236	0.5158	0.5267
Total	0.5680	0.5158	0.5158	0.5158	0.5158

Column 2 above is identical to column 2 in Table 11. Columns 3 through 6 above are the same as column 3 from Tables 7 through 10 respectively.

Table 13
Measures of the Level of Dissimilarity

Distribution of Workers (1)	Index of Dissimilarity (2)	Portion of segregation explained by each method (3)
Actual Distribution	0.4068	-----
Method 1	0.0684	0.1681
Method 2	0.0248	0.0609
Method 3	0.0180	0.0442
Method 4	0.0108	0.0267

Notes: The index of dissimilarity is found as the following

$$S_j = \frac{1}{2}\sum_{i=1}^{4}|M_{ij} - F_{ij}|$$

where M_i is the proportion of the total number of males employed in occupation i and F_i is the proportion of the total number of females employed in occupation i. This measure represents the proportion of male (or female) workers who would have to change jobs for the work-force to be perfectly integrated (Duncan and Duncan, 1955).

Chapter 7

CONCLUSIONS

This study has addressed the issue of gender segregation in the workplace. Specifically, the focus here has been on estimating the significance of individual workers' tastes and preferences on the resulting distribution of those workers across occupations. In other words, this work has focused on occupational gender segregation caused by self-selection. The key to accomplishing this objective has been isolating the effects of tastes on occupational outcome. A model was developed to aid in this process.

This model was designed to reflect a labor market in which only one factor affected occupational outcome, worker's tastes. The first objective in developing the model was to show a relationship between preferences for job characteristics and the job category in which a person works. This was established through an individual's reservation wage for a job. The model developed here demonstrated that an individual's reservation wage for a given job is affected by numerous factors including human capital, net costs of job search, personal characteristics, as well as preferences for the characteristics of the job.

Specifically, these preferences were shown to negatively affect the reservation wage. That is, the more an individual dislikes the characteristics associated with a given occupation, the higher will be their reservation wage for that job. The only assumptions needed to arrive at this result is that search is a sequential process, workers utility is additive in wage and non-wage characteristics, workers search over

numerous job categories, and workers know the distribution of wages for all jobs combined but not for any individual occupation.

After demonstrating how reservation wages are determined and the factors affecting them, the next step was to remove the influence on occupational outcome of non-taste factors affecting the reservation wage (human capital, net search costs, and personal characteristics). This is accomplished by examining reservation wages in several job categories for a single person. This approach holds constant all non-taste characteristics affecting reservation wages since they are individual specific.

Any difference in an individual's reservation wages for different jobs can then be attributed solely to differences in that individual's tastes for the non-wage characteristics for those jobs. When a worker is then matched with one of these jobs, this outcome is not influenced by the non-taste characteristics that help to determine reservation wages. Further, when a worker is matched with a job here, tastes are the only factor affecting the outcome from the supply side.

Having shown this, the next objective was to eliminate any demand side factors that could result in occupational gender segregation. This was accomplished by assuming that employers and prospective employers are completely passive in the labor market. That is, when a prospective employee applies for a job, it is assumed that the employer extends a job offer.

Additionally, the offer extended is at a predetermined wage. This assumption removes many of the factors influencing a worker's occupational outcome from the

demand side, including discrimination. There is, however, a demand side factor that remains even with this assumption.

That factor is the distribution of wages. Assuming that employers are completely passive does not, in and of itself, impose any restrictions on the distribution of wages. A simplifying assumption that wage distributions have zero variance could have been employed. However, in this model wages were allowed to be distributed with a non-zero variance.

To accomplish this in the model, wages were assumed to be distributed randomly by employer. That is, each employer was assumed to offer a specific wage from a given distribution of wages and this offered wage would not change for different applicants. This assumption maintains the passivity of employers that is assumed above since they still have no choice over what wage to offer any applicant.

The assumption that wages are a random variable is not a simplification in this model. In reality, wage is a random variable that has a distribution that can be characterized by a finite mean and variance. By assuming that offered wages in this model were stochastic, the model reflected reality while still meeting the requirement that taste differences be the primary factor affecting occupational outcome. While the distribution of wages does affect the occupation with which a worker is matched, there is no evidence that it contributes to the amount of gender segregation by occupation so there is no significant problem with including it.

The model of the labor market developed herein is based on workers searching for employment. In their search, they encounter potential employers who

make employment offers. The basis for the employee's decision whether or not to accept the offer is based on the reservation wage and the wage offered.

If the offered wage meets or exceeds the individual's reservation wage, the offer is accepted. If the wage offer is below the reservation wage, the employment offer is rejected. A rejected offer is lost forever, and the individual continues their search for a job offer that is acceptable.

When an acceptable offer is found, the worker accepts employment and a match is made. A more complete model of the labor market would also consider separations of workers and employers. This is not necessary for this work, however, since the only interest here is generating a distribution of workers affected primarily by taste differences and examining the amount of predicted gender segregation. This deficiency could be overcome with the simplifying assumption that matches are permanent. However, this gains nothing, so the model of the labor market is left incomplete since the objectives of the study can still be accomplished.

The reservation wages for individuals are then compared to the corresponding wage distributions. The process employed is to compare each of an individual's reservation wages to the corresponding wage distribution and match the individual with a job category based on various criteria. By comparing a point estimate of reservation wage to a distribution of wages, the probability of a worker receiving a wage offer that is acceptable is estimated.

This probability does not, however, capture all of the potential influences on the distribution of workers. Another probability affects the occupation match. This is

the probability of receiving an offer in each job category. The actual process dictates that the worker must first receive the job offer, then determine whether the offered wage exceeds the reservation. While this does introduce another demand side influence on the generated distribution, it seems unlikely that considering the probability of receiving a job offer will introduce any bias into the results.

Accordingly, the probability of receiving a job offer, by job category, is used in estimating the distribution of workers across jobs. The probability of receiving a job offer is estimated via a proxy variable. This variable is the size of the job category relative to all job categories.

Once this is found, the probability of the worker accepting the offered job is estimated using the reservation wage and the distribution of wages for the job category. This latter probability is conditioned on receiving a job offer. This, however, does not affect the method of estimation for the probability of a worker receiving an acceptable offer.

The model developed uses reservation wages across multiple job categories for numerous individuals as well as the distribution of wages in these occupations and other information to generate a distribution of workers. This distribution is determined primarily by workers' tastes. The only significant simplifying assumption needed in this model is that employers are completely passive in terms of their job offers. The next important question was where reservation wages and wage distributions were to be found.

To estimate reservation wages the National Longitudinal Survey, Youth Cohort was used. In 1979 this survey included questions where respondents identified a range in which their reservation wages for different jobs were located. The responses to these questions, as well as other survey questions (intended to capture the effects of the non-taste characteristics on reservation wages), were utilized to arrive at point estimates of reservation wages.

The wage distributions were estimated using the Current Population Survey, March Annual Demographic File data from 1979. Based on the information available in this data set, hourly wages were estimated for respondents who worked in the job categories for which reservation wages are available. Of the respondents in the required job categories, only those who were of comparable age to the respondents for which reservation wage data is available were used to generate the wage distributions. The distributions were estimated such that they were occupation and location specific.

After estimating reservation wages and distributions of wages, workers were matched with occupations. There are several different methods by which a match between a worker and a job can be made. Instead of choosing one method and presenting only these results, several different matching method are used to generate different sets of results.

Each of these different methods of matching workers and jobs was based on the likelihood of a worker receiving an offer and the likelihood of the worker accepting the offer. Stated differently, this latter likelihood is the probability that the

wage offered is above the reservation wage. The only difference between the various methods for matching workers is the amount of information used in estimating these likelihoods.

The first method of matching workers, as well as all other methods, relies on the comparison of reservation wages and estimated wage distributions to estimate the probability than an individual is offered a wage they will accept. In this first method the likelihood of receiving job offers is assumed constant. The results of this method show significant segregation.

The second method used to match workers with jobs improves upon the first by recognizing, and accounting for, the stochastic nature of finding an acceptable wage offer. This is accomplished by matching workers fractionally in each job category. The fraction of a worker placed in each job represents the relative likelihood of a wage offer being received in a job category that is acceptable. The likelihood of receiving job offers is again held constant across occupations. This match results in a distribution of workers that is not significantly segregated.

The last two methods by which workers are matched with jobs offer improvements over the second and third methods, respectively. The specific improvement is that the likelihood of receiving a job offer is allowed to vary across the job categories. These likelihoods are estimated based on the Current Population Survey data referenced previously. The number of persons actually employed in each job category relative to the total number of persons employed in all job categories used here is taken to represent the likelihood of receiving a job offer.

The third method of matching is similar to the first method in that workers are matched with one job category. Due to significant differences in the likelihoods of receiving job offers some job categories cannot be tested for segregation. However, a test for overall segregation results in the finding of no significant gender segregation.

The fourth, and final, method of matching workers improves on the previous method by allowing for fractional matching of workers with jobs. Unlike the results from the third method, tests can be conducted in all of the specific job categories. However, the results from these tests, as well as an overall test, indicate that no significant gender segregation is present.

One interesting outcome from these results is that as the amount of information used in the matching process increases, the significance and amount of segregation decrease. This is easily seen by comparing the results of the first match to the results of the third match. Additionally, by comparing the results of the second and fourth matches, ones sees predicted segregation decrease as information in increased.

Unfortunately, there is no way to test which set of the four different information sets is most valid. There is however, another question along similar lines that is interesting and can be tested. This question is whether it was justified to treat the distributions of wages differently for each job category. By looking at the estimated cumulative density functions of wages in Figures 3 through 6 in Chapter 4, it is clear that each of these estimated density functions are almost perfectly linear. However, this does not tell the full story.

To determine whether the log-wage distributions are statistically different, several tests can be performed. The first logical test is a test for differences in means. When ANOVA is used to test for equality of means for the log-wage distributions, the result is an F-statistic of 42.29. Thus, at even extremely low levels of significance a conclusion that the four job categories used here have different mean wages is justified.

However, differences in means does not, by itself, require that the distributions be estimated separately. If the only difference between the log-wage distributions for the different jobs was average wage, an overall log-wage distribution could be estimated and then shifted differently for each job. This would be very similar to the technique used herein to estimate SMSA specific log-wage distributions. If, however, the variances of the log-wage distributions were different, estimating the density functions specific to occupations would then be required.

Given the data used to construct the log-wage distributions, simple F-tests can be performed to determine whether the distributions constructed have similar variances. These tests are designed to test the hypothesis of equal variances for two normally distributed populations. The variance for each of the log-wage distributions is shown in the second column of Table 14. These variances are calculated from the distributions as computed, prior to the shifting that is made for each SMSA.

A cursory examination of the variance for each distribution reveals they range from 0.1233 to 0.1496, relatively small differences in absolute terms. However, relative to the magnitude of the variances, the range is sizable. The maximum

variance, for the job Cleaning, is 27% greater than the minimum variance, found in the occupation Burgers.

The F-statistic for each test and the level of significance for these values are presented in columns four and five of Table 14 as well. The F-statistic for these tests is merely the ratio of the two variances being tested, constructed in such a way that the result has a value of at least one. This restriction on the F-statistic is necessary since it is designed as a one-tailed test.

All of the tests result in a failure to reject the null hypothesis that the distributions have equal variances at any commonly accepted level of significance. These tests then leave the impression that the results from the first method of matching workers may be valid. However, this conclusion may not be warranted.

The F-test for equal variances is designed for distributions that are in the family of normal distributions. It is unknown at this point whether the log-wage distributions used here are distributed normally. In order to test for this, higher order moments for the wage distributions are needed. The first through fourth moments for the four log-wage distributions are presented in Table 15. Based on these moments, estimates of the skewness and kurtosis of each distribution can be found.

These estimates are presented in Table 16. Also present in this table is the Jarque-Bera (JB) statistic. This statistic is designed to test the assumption of a variable following a normal distribution (Gujarati, 1995). This statistic has a distribution that is Chi-squared with two degrees of freedom. The final column in Table 16 indicates the level of significance with which the null hypothesis of the

distribution being normal can be rejected. Clearly, this hypothesis can be rejected in every job category at any conventional level of significance.

Thus, the findings from Table 16 invalidate the results of the hypothesis tests detailed in Table 15. Unfortunately, no information has been gained. It is equally unfortunate that no statistical test exists that would allow for testing of the variances of the wage distributions since wages have been shown to be non-normally distributed. Based on the results of the tests of the log-wage distributions, it is unclear whether estimation of the density functions must be occupation specific. Thus, it seems logical to err on the side of caution and form the density estimates specific to each job category as was done here.

This work has contributed to the already vast economic literature in the area of discrimination and segregation analysis. Several studies addressing segregation have had self-selection as their primary focus. This study, however, is unique in that it is the only one to actually create a distribution of workers across jobs that is determined almost solely by preferences. Previous works in this area have, at most, estimated differences in probabilities of working in certain job categories between males and females, which can be attributed to self-selection.

Similarly to most previous works, this study finds some evidence that taste differences by gender for non-pecuniary work characteristics do affect occupational outcomes. However, unlike most previous work on the impact of self-selection, these results are by no means overwhelming. In only one of four sets of results presented herein can the hypothesis of independence between gender and occupation be

rejected. Further, the results of this study do not seem to explain the pattern of occupational gender segregation observed in the actual labor market.

The findings of this study indicate that self-selection is not a major factor causing gender segregation in the workplace. This information may influence policy designed to address discrimination that is manifest through segregation. If none of the segregation observed in a particular industry or occupation is caused by differences in worker's preferences for the non-pecuniary characteristics of the jobs, then public policy would need to set its target at a lower level of segregation than in a case where self-selection does cause segregation. Even if self-selection does not cause segregation, segregation may exist that is caused by other supply side factors such as difference in human capital investment decisions. This is why a lower, but possibly non-zero, target would be set if self-selection were not a factor.

Having found such results, a worthy exercise at this point is to discuss some of the weaknesses of this work and specifically how these possibly can be overcome in future research. Several weaknesses are a function of the data used to estimate reservation wages. Namely, the relatively small number of job categories, the type of jobs used, the age of respondents, and the lack of timeliness of the data.

Only having four job categories over which to distribute workers causes some degree of concern since it is in no way a realistic assumption that only four job categories exist in the labor market. Additionally, if only four job categories were to be used in a study of this sort, it would seem prudent to select jobs that better

represent the overall economy. However, choosing different jobs could present other problems not encountered here.

For example, if non-entry level jobs such as engineer, architect, or teacher were used in such a study, many questions would arise about the responses. It would be unclear how to interpret the results of such questions since there are differences in these jobs other than just the non-pecuniary characteristics. If a person were asked at what wage they would accept a job as an engineer, they would surely consider the costs associated with acquiring the necessary human capital to work as an engineer in forming their answer. These effects would not then present a clear picture of differences in reservation wages caused purely by differences in the characteristics associated with different jobs.

The only way to ensure that differences in reservation wages are caused strictly by preferences for the non-wage characteristics of the different jobs is to use only occupations in which all survey respondents could immediately work. That is, only jobs that would require virtually no additional human capital investment could be used in a survey of respondents that represent the potential workforce. The results from the questions about willingness to accept employment in such jobs would provide a true indication of the individual's perception of the characteristics of those jobs.

This, in fact, is the approach used in the NLSY. The list of jobs that are included in a survey should be jobs that respondents in the sample would see as their relevant set of immediate employment opportunities. A list of these jobs would then

depend on the universe of sample respondents. For the NLSY in 1979, the relevant universe was individuals between the ages of 14 and 22.

This is another possible weakness of the data employed to estimate reservation wages, the age of survey respondents. For the purposes of this study, only those individuals between the ages of 16 and 22 were utilized. A survey that was not restricted to only those under 22 years of age would certainly provide a more representative view of the working age population. This lack of representativeness may be a reason for the lack of correlation between the results of this work and observed gender segregation patterns. If the taste differences between genders are less pronounced for youths than for the overall labor force, then one should expect to find less gender segregation caused by self-selection among younger workers.

Both of these weaknesses, few jobs and restricted sample, could be overcome in a new survey. Another weakness of the data from the 1979 NLSY is the age. This data is nearly twenty years old. While it does not face degradation due to its age, it certainly would be preferable to have data that is more recent. While this would not necessarily provide better information it would allow for certain updating that could be beneficial. This could be especially important in the public policy arena. Making public policy based on data that is nearly 20 years old could lead to serious policy errors.

The ideal survey for this type of work would be one that sampled individuals at least 16 years of age in such a way as to represent the potential workforce. It would ask the series of questions about accepting job offers for more job categories

than the NLSY. Possible job categories that would provide useful information include telemarketer, stockperson at a grocery or discount store, landscape or grounds maintenance, day care worker, data entry operator, secretary, receptionist, convenience store clerk, cook, waitperson, library clerk, bank teller, parking attendant, amusement park worker, garbage collector, cafeteria worker, as well as the four job categories used herein. A respondent could begin work in any of these occupations with little or no investment in human capital. Accordingly, the differences in reservation wages between any two jobs could be attributed solely to tastes for the non-pecuniary characteristics for those jobs.

A final weakness in this study that results from the data used to estimate reservation wages is the job categories. While the job categories are extremely specific, in some cases the categories do not correspond well with occupations in the Current Population Survey. This is especially true with the occupations Cleaning and Supermarket. Both of these categories were matched with a closely corresponding occupation in the CPS, but the matches were not as exact as with Burger and Dishes. A future survey would ideally utilize job categories from the CPS, or some other such survey, in defining reservation wage questions.

Another set of weaknesses arises from the use of CPS data to construct the wage distributions. First, a wage estimate that is not dependent on the numerous assumptions used here would be preferable. The wage estimate is formed based on the assumptions that the respondent worked all of the prior year in only one job category and that all of the income earned was from work in that job.

The estimate was calculated by dividing the amount of income received from work in the prior year by the product of the average hours worked per week and the number of hours worked in the prior year. Even slight misrepresentations in any of these three responses could result in nontrivial errors in the estimated hourly wage. A better method of collecting the necessary information would be to ask survey respondents the hourly wage received for each job held in the prior year.

Second, the wage distributions would better suit this study if they were both occupation and location specific. That is, instead of estimating by how much average wages are different by location, distributions would, ideally, be estimated based on observations specific to that location. As mentioned in the text, ideal wage distributions would be based on observations from the specified job category in the specific location. This would require a much larger sample in the CPS that was used in 1979.

There are also several weaknesses that are based on the implementation of the model. First among these is the level of detail on the respondent's location. The identifier used in this study is SMSA. However, a given SMSA typically covers numerous counties in an area. While labor market conditions in bordering counties are likely to be similar, they may not be identical. In fact, there may be radical differences in conditions across the SMSA.

Consider the SMSA 8840, Washington, D.C. - Maryland - Virginia. This SMSA covers counties in two different states and the District of Columbia. While the

labor markets in the three different areas may be similar, there remains a high likelihood of non-trivial differences.

For example, significant differences could be introduced by differences in tax rates in the three different areas. Additionally, several counties in the SMSA have substantial rural components. It seems quite likely that these rural areas would have a drastically different labor market than that in the District of Columbia.

Because of these, and other, problems with using SMSA, some other identifier of location would offer an improvement. Using counties would provide better information about the labor market faced by the specific individual. However, this would not eliminate the problems. Many counties have both urban and rural components, especially in western states with large counties. Thus, a smaller geographic identifier may be preferable to county.

The next logical step down is to the city. However, this may not be desirable due to the fact that many people chose to live in one city and work in a different city. For that matter, many people chose to live and work in different counties.

Based on these possibilities, there is no clear answer as to what level of detail should be employed in determining the location of a respondent. Ideally, the researcher would have some knowledge of the area in which the respondent resided and could then make some determination as to what level of detail should be employed. This, however, would be an extremely costly approach to such a study.

Another problem in this study is the estimate of the probability of a worker receiving a job offer. This is estimated by the relative number of persons employed

in each job. This ignores other possible influences on this likelihood. The probability of a worker receiving an employment offer in a certain job category is a function of at least two things.

First, the worker must find an employer that employs workers in the specified occupation. Second, the employer must have a job opening. The former influence is captured by the estimate of the probability of receiving a job offer in a given job category that is employed. The latter, however, is not considered here.

There are several factors that affect the likelihood of an employer of a specific job category having an opening. The most important determinant of this probability will be the rate at which either employees or the employer terminate the match, commonly referred to as the turnover rate. The turnover rate is affected by numerous factors including wages and non-wage characteristics.

For this work, it was assumed that matches between employers and employees are permanent. Accordingly, turnover is not possible in this model and was therefore ignored. Since turnover is ignored in this model, possible variation in turnover rates by occupation are also not considered.

It is quite likely, however, that the turnover rates will vary from one job category to the next. Accordingly, this variation should be considered. An estimate of the turnover rate for each job category used should thus be obtained. The turnover rate would indicate the probability of the existence of an opening.

To see this consider the following example, suppose the turnover rate in a certain job is 20% per year. This indicates that 20% of the jobs in this occupation

were vacant during some portion of the year. Thus, the probability of an employer having an opening in this job category during the year is 20%.

Once job specific turnover rates are obtained, the probability of a worker finding a job offer could be estimated as the product of the probability of a worker finding an employer of the job category and the probability of such an employer having a vacancy. This, of course, assumes independence between these two probabilities. Allowing turnover rates to affect the match of employees and employers, and thus affect the distribution of workers, should not introduce any problems in this model since employers can still be assumed passive in the matching process. The employer simply offers a job if they have a vacancy, based on the job specific turnover rate.

A final potential concern about this work that is caused by the methods employed is the forced matching of every employee to a job. That is, the respondents in the NLSY are all 'forced' into one of the four job categories. This is by no means an accurate representation of a well functioning labor market. In fact, many of the respondents in the samples identified themselves as out of the labor force. A good question is whether these respondents, as well as others, should be forced into one of the four job categories.

The easiest argument in favor of forced matching is that these matches are in fact not forced. In this model, which represents an artificial labor market in which all respondents are potential workers, respondents are not 'forced' to accept jobs. They

still use their reservation wage as the determining factor in their decision whether to accept or reject an employment offer.

Accordingly, workers will only accept employment if doing so maximizes their utility. The appropriate assumption in this model is not that workers are forced into one of the four jobs used. The proper set of assumptions is that all respondents search for, and find, an acceptable offer in one of these four occupations. These are, in fact, the assumptions that are made.

Additionally, this model is intended to estimate a distribution of workers that is determined primarily by taste differences. Allowing for respondents to not be matched with one of the four jobs would result in non-taste factors influencing the estimated distribution of workers. To see this consider the possibility of omitting workers from the sample based on their reservation wage. Since individuals with high reservation wages are more likely to not search for work, these are the respondents that should be left out.

However, it is quite likely that an individual's reservation wage is high because they have a great deal of human capital. Thus such a practice would allow human capital to influence the distribution of workers. Accordingly, omitting workers from the sample is not a wise choice. The only alternative is to assume that workers do, in fact, search for employment.

Clearly this study can be improved upon in future research. Many of the problems with this work are related to the data used. Unfortunately, finding data that is ideally suited for any given research project is often times a difficult task. For the

questions that arise from the model developed and the approach used here, there are no definitive answers indicating how these questions should be addressed in future efforts.

This work has improved on the existing literature that examines gender segregation in the workplace in two ways. First, a distribution of workers is estimated. Existing research in this area has, at best, estimated differences in the probabilities of individuals working in different job categories. Second, the generated distribution of workers across job categories is constructed in such a way that it is unaffected by any factors other than taste differences and the distribution of wages in the different occupations.

The approach developed here has applications in areas other than just the study of labor market segregation. This technique can be used to examine any market where segregation is observed. One of the most readily available applications is in the housing market. The segregation observed in the housing market tends to follow racial lines. The question that this approach could address in this instance is estimating what portion of the observed housing segregation is caused by self-selection.

Another area in which this technique could be applied is markets where discrimination is thought to be present. For example, an extension of the work presented here could examine the expected wage for the workers when they are matched with jobs. If any differences exist in the average expected wage of different groups of workers, these differences could be attributed to self-selection. Other

possible applications of this method include the markets for insurance or home-financing.

The results of the current work are mixed, with some distributions of workers across occupations showing very little impact of self-selection and others indicating that gender and occupational outcome are related. On net, however, the results from the data used here imply that gender differences in tastes have very little impact on the occupational outcome of workers. Further, these gender based taste differences for non-pecuniary work characteristics seem to not be able to explain observed patterns of gender segregation in the workplace since many occupations are predicted to be male (female) dominated when, in reality, that job category is female (male) dominated.

These results, however, are by no means conclusive. In fact, applied strictly, this study finds that on average self-selection does not affect the occupation choice of relatively young workers when they are choosing over only four occupations which can all be classified as low-wage entry level positions. This then implies that the distribution of these young workers across the specific occupations examined is not influenced by self-selection. To claim that such results holds throughout the labor market would be ill-advised.

Table 14
Test for Equality of Variances of Wage Distributions

Job Category (1)	Variance (2)	Variance Ratio (3)	F- Statistic (4)	Level of Significance (5)
Burgers	0.1233	Cleaning / Burgers	1.27275	0.1542
Cleaning	0.1496	Cleaning / Dishes	1.19824	0.3769
Dishes	0.1268	Cleaning / Market	1.03337	0.7372
Market	0.1481	Dishes / Burgers	1.06218	0.7853
		Market / Burgers	1.23165	0.2360
		Market / Dishes	1.15955	0.4843

Table 15
Moments of Log-Wage Distributions

Job Category (1)	Observations (2)	Mean (3)	Variance (4)	Third Moment (5)	Fourth Moment (6)
Burgers	90	1.3924	0.1233	0.03949	0.04557
Cleaning	612	1.5914	0.1496	0.03105	0.06294
Dishes	62	1.4259	0.1268	0.0581	0.06252
Market	346	1.5213	0.1481	0.06244	0.09104
Total	1110	1.5442	0.1494	0.04371	0.07156

Variance and higher order moments are defined as:

$$\sigma^i(X) = E(X - \mu)^i \quad \forall\, i \geq 2$$

These values are estimated by the function:

$$s^i(X_j) = \frac{\sum_{j=1}^{n} (X_j - \overline{X})^i}{n - 1} \quad \forall\, i \geq 2$$

Table 16
Measures of Skewness, Kurtosis, and Test of Normal Distribution

Job Category (1)	Skewness (2)	Kurtosis (3)	Jarque-Bera Statistic (4)	Level of Significance (5)
Burgers	0.8325	2.9989	70.6873	0.005530
Cleaning	0.2880	2.8126	9.3557	0.009299
Dishes	1.2656	3.8871	183.4523	$9.2111 * 10^{-5}$
Market	1.2000	4.1501	180.6015	$6.7333 * 10^{-23}$

Skewness is defined as:

$$S = \frac{\left[E(X - \mu)^3\right]^2}{\left[E(X - \mu)^2\right]^3}$$

Kurtosis is defined as:

$$K = \frac{E(X - \mu)^4}{\left[E(X - \mu)^2\right]^2}$$

The Jarque-Bera (JB) statistic is designed to test the distribution of a random variable to determine if it is normally distributed. The JB statistic has a Chi-squared distribution with 2 degrees of freedom. It is defined as:

$$JB = n\left[\frac{S^2}{6} + \frac{(K - 3)^2}{24}\right]$$

Bibliography

Aigner, Dennis and Cain, Glen, " Statistical Theories of Discrimination in the Labor Market", *Industrial and Labor Relations Review*, 1977, 30, 175 – 187.

Arrow, Kenneth, "The Theory of Discrimination", in *Discrimination in Labor Markets*, O. Ashenfelter and A. Rees, eds., Princeton, NJ, Princeton University Press, 1973, 3 – 33.

Blau, Francine and Ferber, Marianne, *The Economics of Women, Men, and Work*, 2 ed., New Jersey, Prentice – Hall, 1992.

Blau, Francine and Hendricks, Wallace, "Occupational Segregation by Sex: Trends and Prospects", *American Economic Review*, 1979, 12(2), 316 – 320.

Becker, Gary, *The Economics of Discrimination*, 1973, Chicago, The University of Chicago Press, (Original Edition, 1957).

Center for Human Resource Research, 1994

Daymont, Thomas and Andrisani, Thomas, "Job Preferences, College Major, and the Gender Gap in Earnings", *The Journal of Human Resources*, 1984, 19, 408 – 428.

Devine, Theresa and Kiefer, Nicholas, *Empirical Labor Economics: The Search Approach*, 1991, New York, Oxford, University Press.

Duncan, Otis and Duncan, Beverly, "A Methodological Analysis of Segregation Indexes", *American Sociological Review*, 1955, 20, 210 – 218.

England, Paula, "The Failure of Human Capital Theory to Explain Occupational Sex Segregation", *The Journal of Human Resources*, 1982, 17(3), 358 – 370.

Feldstein, Martin and Poterba, James, "Unemployment Insurance and Reservation Wages", *Journal of Public Economics*, 1981, 23, 141 – 167.

Filer, Randall, "The Role of Personality and Tastes in Determining Occupational Structure", *Industrial and Labor Relations Review*, 1986, 39, 412 – 424.

Filer, Randall, "Occupational Segregation, Compensating Differentials, and Comparable Worth", in *Pay Equity: Empirical Inquiries*, Michael, Robert, Hartmann, Heidi, and O'Farrell, Brigidet, eds. Washington, D.C.: National Academy Press, 1989.

Gastwirth, Joseph, "On Probabilistic Models of Consumer Search for Information", *Quarterly Journal of Economics*, 1976, 90, 38 – 50.

Gujurati, Damodar, *Basic Econometrics*, Third ed., New York, McGraw – Hill, Inc., 1995.

Gupta, Nabatina, "Probabilities of Job Choice and Employer Selection and Male – Female Occupational Differences", *American Economic Review*, 193, 83(2), 57 – 61.

Hofler, Richard and Murphy, Kevin, "Estimating Reservation Wages of Employed Workers Using a Stochastic Frontier", *Southern Economic Journal*, 1994, 60(4), 961 – 976.

Jensen, Peter and Westergård – Nielsen, Niels, "A Search Model Applied to the Transition from Education to Work", *Review of Economic Studies*, 1987, 52, 461 – 472.

Khandker, Rezaul, "Offer Heterogeneity in a Two State Model of Sequential Search", *The Review of Economics and Statistics*, 1988, 70, 259 – 265.

Kiefer, Nicholas and Neumann, George, "An Empirical Job – Search Model, With a Test of the Constant Reservation – Wage Hypothesis", *Journal of Political Economy*, 1979, 87(1), 89 – 107.

Leonard, Justin, "Women and Affirmative Action", *Journal of Economic Perspectives*, 1989, 3(1), 61 – 75.

Manning, Richard and Morgan, Peter, "Search and Consumer Theory", *Review of Economic Studies*, 1982, 49, 203-216.

McCall, J.J., "Economics of Information and Job Search", *Quarterly Journal of Economics*, 1970, 85, 111 – 126.

McCue, Kristin and Reed, W. Robert, "New Evidence on Workers' Willingness to Pay for Job Attributes", *Southern Economic Journal*; 1996,62(3), 627 – 652.

Michael, Robert, and Hartmann, Heidi, "Pay Equity: Assessing the Issues" in *Pay Equity: Empirical Inquiries*, Michael, Robert, Hartmann, Heidi, and O'Farrell, Brigidet, eds. Washington, D.C.: National Academy Press, 1989.

Morgan, Peter and Manning, Richard, "Optimal Search", *Econometrica*, 1985, 53, 923 – 944.

Mortensen, Dale, "Job Search and Labor Market Analysis", in *Handbook of Labor Economics*, Vol II, Ashenfelter, Orley and Layard, Richard, eds., Amsterdam, North – Holland, 848 – 919.

Niesing, Willem, van Praag, Bernard, and Veenman, Justus, "The Unemployment of Ethnic Minority Groups in the Netherlands", *Journal of Econometrics*, 1994, 61(1), 173 - 196.

Pencavel, John, "Labor Supply of Men: A Survey", ", in *Handbook of Labor Economics*, Vol I, Ashenfelter, Orley and Layard, Richard, eds., Amsterdam, North – Holland, 3 102.

Phelps, Edmund, "The Statistical Theory of Racism and Sexism", *American Economic Review*, 1972, 62, 659 – 666.

Polachek, Solomon, "Occupational Segregation: An Alternative Hypothesis", *International Economic Review*, 1975, 16, 451 – 470.

Polachek, Solomon, "Occupational Self-Selection: A Human Capital Approach to Sex Differences in Occupational Structure", *The Review of Economics and Statistics*, 1981, 63(1), 60 – 69.

Reskin, Barbara and Hartmann, Heidi, "Explaining Sex Segregation in the Workplace", in *Women's Work, Men's Work: Sex Segregation on the Job*, Reskin, Barbara and Hartmann, Heidi, eds., Washington, D.C.: National Academy Press, 1986.

Rothschild, Michael, "Searching for the Lowest Price When the Distribution of Prices is Unknown", *Journal of Political Economy*, 1974, 82, 689 - 711.

Stigler, George, "The Economics of Information", *Journal of Political Economy*, 1961, 69, 213 – 225.

Stigler, George, "Information in the Labor Market", *Journal of Political Economy*, 1962, 70, 94 – 105.

United States Bureau of the Census, *Current Population Survey, March 1979 Technical Documentation*, 1980, Washington, D.C., Bureau of the Census.

United States Department of Labor, *Employment and Earnings*, 1990, Washington, D.C., Department of Labor.

New Webster's Dictionary and Thesaurus of the English Language, 1992. Danbury, CT, Lexicon Publications.

Appendix A
Estimation of Reservation Wages for Male Sample in Occupation Burgers

```
/*        M_BURGER.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*Opening file containing data.                    */

open f1 = h:\thesis\data\data1\male\male.dat for read;
x = readr(f1,1441);

/*      Specifying data matrices, x1 contains independent variables, y1 – y1 contain
identifies for the self-reported reservation wage category.          */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*      The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.          */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xb1 = x1~d2~d3~d4~c2~c3~c4~s2~s3~s4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y4,x1);
b4 = olsqr(y4,xb1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b4 = b4|0.5;

__title = "Logistic estimation of Burgers for Males";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(g1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
g2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
g3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
g4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 1;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.          */

output file=bbeta_m.out reset;

{bbeta1,f1,ga1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{bbeta2,f2,ga2,h2,retc2} = maxprt(maxlik(xb1,0,&lli,b4));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{bbeta3,f3,ga3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1-(f2/f3);
print "Psuedo R-square =" prsq;

output off;

/*      Approximationg the error terms on the reservation wage equation.          */

bx = rows(xb1);
rb = rows(bbeta2);
vb = seqa(1,1,rb-1)';
bbeta4 = submat(bbeta2,vb,0);
bhat = xb1*bbeta4;

b = (maxc(0*g1'|g2'|2*g3'|3*g4'))';
bi1 = zeros(2,bx);
```

```
i = 1;
j = 1;

do while i le bx;
   if b[.,i] == 0; bi1[2,i] = (((0.1)-bhat[i,.])/bbeta2[rb,.]);
elseif b[.,i] == 1; bi1[2,i] = ((ln(2.50)-bhat[i,.])/bbeta2[rb,.]);
elseif b[.,i] == 2; bi1[2,i] = ((ln(3.50)-bhat[i,.])/bbeta2[rb,.]);
          else; bi1[2,i] = ((ln(5.00)-bhat[i,.])/bbeta2[rb,.]);
   endif;
   i = i+1;
endo;

do while j le bx;
   if b[.,j] == 0; bi1[1,j] = ((ln(2.50)-bhat[j,.])/bbeta2[rb,.]);
elseif b[.,j] == 1; bi1[1,j] = ((ln(3.50)-bhat[j,.])/bbeta2[rb,.]);
elseif b[.,j] == 2; bi1[1,j] = ((ln(5.00)-bhat[j,.])/bbeta2[rb,.]);
          else; bi1[1,j] = ((ln(15)-bhat[j,.])/bbeta2[rb,.]);
   endif;
   j = j+1;
endo;

proc integr_b(eps);
   local rb;
   rb = rows(bbeta2);
   retp((eps.*bbeta2[rb,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

bint = intquad1(&integr_b,bi1);
_intord = 40;

bint1=(1./(1+exp(-bi1)));
bint2=bint1[1,.]-bint1[2,.];

bint3=bint./bint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_b = bhat+bint3;

v10 = 39;
smsa = submat(x,0,v10);

output file=m_burger.out reset;
```

```
print lnrswg_b~smsa;

output off;
```

Appendix B
Estimation of Reservation Wages for Female Sample in Occupation Burgers

```
/*        F_BURGER.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containing data.        */

open f1 = h:\thesis\data\data1\female\female.dat for read;
x = readr(f1,1535);

/*        Specifying data matrices, x1 contains independent variables, y1 – y4 contain
identifiers for the self-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xb1 = x1~d2~d3~d4~c2~c3~c4~s2~s3~s4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y4,x1);
b4 = olsqr(y4,xb1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b4 = b4|0.5;

__title = "Logistic estimation of Burgers for Females";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(g1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
g2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
g3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
g4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 0;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.          */

output file=bbeta_f.out on;

{bbeta1,f1,ga1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{bbeta2,f2,ga2,h2,retc2} = maxprt(maxlik(xb1,0,&lli,b4));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{bbeta3,f3,ga3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-square =" prsq;

output off;

/*      Approximating the error terms on the reservation wage equation.          */

bx = rows(xb1);
rb = rows(bbeta2);
vb = seqa(1,1,rb-1)';
bbeta4 = submat(bbeta2,vb,0);
bhat = xb1*bbeta4;

b = (maxc(0*g1'|g2'|2*g3'|3*g4'))';
bi1 = zeros(2,bx);
```

```
i = 1;
j = 1;

do while i le bx;
   if b[.,i] == 0; bi1[2,i] = (((0.1)-bhat[i,.])/bbeta2[rb,.]);
elseif b[.,i] == 1; bi1[2,i] = ((ln(2.50)-bhat[i,.])/bbeta2[rb,.]);
elseif b[.,i] == 2; bi1[2,i] = ((ln(3.50)-bhat[i,.])/bbeta2[rb,.]);
          else; bi1[2,i] = ((ln(5.00)-bhat[i,.])/bbeta2[rb,.]);
   endif;
   i = i+1;
endo;

do while j le bx;
   if b[.,j] == 0; bi1[1,j] = ((ln(2.50)-bhat[j,.])/bbeta2[rb,.]);
elseif b[.,j] == 1; bi1[1,j] = ((ln(3.50)-bhat[j,.])/bbeta2[rb,.]);
elseif b[.,j] == 2; bi1[1,j] = ((ln(5.00)-bhat[j,.])/bbeta2[rb,.]);
          else; bi1[1,j] = ((ln(15)-bhat[j,.])/bbeta2[rb,.]);
   endif;
   j = j+1;
endo;

proc integr_b(eps);
   local rb;
   rb = rows(bbeta2);
   retp((eps.*bbeta2[rb,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

bint = intquad1(&integr_b,bi1);
_intord = 40;

bint1=(1./(1+exp(-bi1)));
bint2=bint1[1,.]-bint1[2,.];

bint3=bint./bint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_b = bhat+bint3;

v10 = 39;
smsa = submat(x,0,v10);

output file=f_burger.out reset;
```

```
print lnrswg_b~smsa;

output off;
```

Appendix C
Estimation of Reservation Wages for Female Sample in Occupation Cleaning

```
/*        M_CLEAN.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containing data.        */

open f1 = h:\thesis\data\data1\male\male.dat for read;
x = readr(f1,1441);

/*        Specifying data matrices, x1 contains independent variables, y1 – y4 contain
identifiers for the slef-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xc1 = x1~d2~d3~d4~s2~s3~s4~g2~g3~g4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y2,x1);
b2 = olsqr(y2,xc1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b2 = b2|0.5;

__title = "Logistic estimation of Cleaning for Males";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(c1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
c2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
c3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
c4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 0;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.         */

output file=cbeta_m.out on;

{cbeta1,f1,g1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{cbeta2,f2,g2,h2,retc2} = maxprt(maxlik(xc1,0,&lli,b2));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{cbeta3,f3,g3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-square =" prsq;

output off;

/*      Approximating the error terms on the reservation wage equations.          */

cx = rows(xc1);
rc = rows(cbeta2);
vc = seqa(1,1,rc-1)';
cbeta4 = submat(cbeta2,vc,0);
chat = xc1*cbeta4;

c = (maxc(0*c1'|1*c2'|2*c3'|3*c4'))';
ci1 = zeros(2,cx);
```

```
i = 1;
j = 1;

do while i le cx;
   if c[.,i] == 0; ci1[2,i] = (((0.1)-chat[i,.])/cbeta2[rc,.]);
elseif c[.,i] == 1; ci1[2,i] = ((ln(2.50)-chat[i,.])/cbeta2[rc,.]);
elseif c[.,i] == 2; ci1[2,i] = ((ln(3.50)-chat[i,.])/cbeta2[rc,.]);
          else; ci1[2,i] = ((ln(5.00)-chat[i,.])/cbeta2[rc,.]);
   endif;
   i = i+1;
endo;

do while j le cx;
   if c[.,j] == 0; ci1[1,j] = ((ln(2.50)-chat[j,.])/cbeta2[rc,.]);
elseif c[.,j] == 1; ci1[1,j] = ((ln(3.50)-chat[j,.])/cbeta2[rc,.]);
elseif c[.,j] == 2; ci1[1,j] = ((ln(5.00)-chat[j,.])/cbeta2[rc,.]);
          else; ci1[1,j] = ((ln(15)-chat[j,.])/cbeta2[rc,.]);
   endif;
   j = j+1;
endo;

proc integr_c(eps);
   local rc;
   rc = rows(cbeta2):
   retp((eps.*cbeta2[rc,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

c_int = intquad1(&integr_c,ci1);
_intord = 40;

cint1=(1./(1+exp(-ci1)));
cint2=cint1[1,.]-cint1[2,.];

cint3=c_int./cint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_c = chat+cint3;

output file=m_clean.out reset;

print lnrswg_c;

output off;
```

Appendix D
Estimation of Reservation Wages for Female Sample in Occupation Cleaning

```
/*        F_CLEAN.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containing data.        */

open f1 = h:\thesis\data\data1\female\female.dat for read;
x = readr(f1,1535);

/*        Specifying data matrices. x1 contains independent variables, y1 – y4 contain
identifiers for the self-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xc1 = x1~d2~d3~d4~s2~s3~s4~g2~g3~g4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y2,x1);
b2 = olsqr(y2,xc1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b2 = b2|0.5;

__title = "Logistic estimation of Cleaning for Females";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(c1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
c2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
c3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
c4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 1;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.          */

output file=cbeta_f.out reset;

{cbeta1,f1,g1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{cbeta2,f2,g2,h2,retc2} = maxprt(maxlik(xc1,0,&lli,b2));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{cbeta3,f3,g3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-square =" prsq;

output off;

/*      Approximating the error terms on the reservation wage equation.          */

cx = rows(xc1);
rc = rows(cbeta2);
vc = seqa(1,1,rc-1)';
cbeta4 = submat(cbeta2,vc,0);
chat = xc1*cbeta4;

c = (maxc(0*c1'|c2'|2*c3'|3*c4'))';
ci1 = zeros(2,cx);
```

```
i = 1;
j = 1;

do while i le cx;
   if c[.,i] == 0; ci1[2,i] = (((0.1)-chat[i,.])/cbeta2[rc,.]);
elseif c[.,i] == 1; ci1[2,i] = ((ln(2.50)-chat[i,.])/cbeta2[rc,.]);
elseif c[.,i] == 2; ci1[2,i] = ((ln(3.50)-chat[i,.])/cbeta2[rc,.]);
          else; ci1[2,i] = ((ln(5.00)-chat[i,.])/cbeta2[rc,.]);
   endif;
   i = i+1;
endo;

do while j le cx;
   if c[.,j] == 0; ci1[1,j] = ((ln(2.50)-chat[j,.])/cbeta2[rc,.]);
elseif c[.,j] == 1; ci1[1,j] = ((ln(3.50)-chat[j,.])/cbeta2[rc,.]);
elseif c[.,j] == 2; ci1[1,j] = ((ln(5.00)-chat[j,.])/cbeta2[rc,.]);
          else; ci1[1,j] = ((ln(15)-chat[j,.])/cbeta2[rc,.]);
   endif;
   j = j+1;
endo;

proc integr_c(eps);
   local rc;
   rc = rows(cbeta2);
   retp((eps.*cbeta2[rc,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

c_int = intquad1(&integr_c,ci1);
_intord = 40;

cint1=(1./(1+exp(-ci1)));
cint2=cint1[1,.]-cint1[2,.];

cint3=c_int./cint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_c = chat+cint3;

output file=f_clean.out reset;

print lnrswg_c;

output off;
```

Appendix E
Estimation of Reservation Wages for Male Sample in Occupation Dishes

```
/*        M_DISHES.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containing data.        */

open f1 = h:\thesis\data\male\male.dat for read;
x = readr(f1,1441);

/*        Specifying data matrices, x1 contains independent variables, y1 – y4 contain
identifiers for the self-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xd1 = x1~c2~c3~c4~s2~s3~s4~g2~g3~g4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y1,x1);
b1 = olsqr(y1,xd1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b1 = b1|0.5;

__title = "Logistic estimation of Dishes for Males";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(d1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
d2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
d3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
d4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t)))):
endp;

/*      Specifying the settings for the optimization process.      */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 1;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.      */

output file=dbeta_m.out reset;

{dbeta1,f1,g1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{dbeta2,f2,g2,h2,retc2} = maxprt(maxlik(xd1,0,&lli,b1));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{dbeta3,f3,g3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-squared =" prsq;

output off;

/*      Approximating the error terms on the reservation wage equation.      */

dx = rows(xd1);
rd = rows(dbeta2);
vd = seqa(1,1,rd-1)';
dbeta4 = submat(dbeta2,vd,0);
dhat = xd1*dbeta4;

d = (maxc(0*d1'|1*d2'|2*d3'|3*d4'))';
di1 = zeros(2,dx);
```

```
i = 1;
j = 1;

do while i le dx;
   if d[.,i] == 0; di1[1,i] = ((0.1-dhat[i,.])/dbeta2[rd,.]);
elseif d[.,i] == 1; di1[1,i] = ((ln(2.50)-dhat[i,.])/dbeta2[rd,.]);
elseif d[.,i] == 2; di1[1,i] = ((ln(3.50)-dhat[i,.])/dbeta2[rd,.]);
          else; di1[1,i] = ((ln(5.00)-dhat[i,.])/dbeta2[rd,.]);
   endif;
   i = i+1;
endo;

do while j le dx;
   if d[.,j] == 0; di1[2,j] = ((ln(2.50)-dhat[j,.])/dbeta2[rd,.]);
elseif d[.,j] == 1; di1[2,j] = ((ln(3.50)-dhat[j,.])/dbeta2[rd,.]);
elseif d[.,j] == 2; di1[2,j] = ((ln(5.00)-dhat[j,.])/dbeta2[rd,.]);
          else; di1[2,j] = ((ln(15)-dhat[j,.])/dbeta2[rd,.]);
   endif;
   j = j+1;
endo;

proc integr_d(eps);
   local rd;
   rd = rows(dbeta2);
   retp((eps.*dbeta2[rd,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

dint = intquad1(&integr_d,di1);
_intord = 40;

dint1=(1./(1+exp(-di1)));
dint2=dint1[1,.]-dint1[2,.];

dint3=dint./dint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_d = dhat+dint3;

output file=m_dishes.out reset;

print lnrswg_d;

output off;
```

Appendix F
Estimation of Reservation Wages for Female Sample in Occupation Dishes

```
/*        F_DISHES.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containing data.        */

open f1 = h:\thesis\data\female\female.dat for read;
x = readr(f1,1535);

/*        Specifying data matrices, x1 contains independent variables, y1 – y4 contain
identifiers for the self-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xd1 = x1~c2~c3~c4~s2~s3~s4~g2~g3~g4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y1,x1);
b1 = olsqr(y1,xd1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b1 = b1|0.5;

__title = "Logistic estimation of Dishes for Females";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(d1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
d2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
d3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
d4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 1;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.          */

output file=dbeta_f.out reset;

{dbeta1,f1,g1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{dbeta2,f2,g2,h2,retc2} = maxprt(maxlik(xd1,0,&lli,b1));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{dbeta3,f3,g3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-square =" prsq;

output off;

/*      Approximating the error terms on the reservation wage equation.          */

dx = rows(xd1);
rd = rows(dbeta2);
vd = seqa(1,1,rd-1)';
dbeta4 = submat(dbeta2,vd,0);
dhat = xd1*dbeta4;

d = (maxc(0*d1'|1*d2'|2*d3'|3*d4'))';
di1 = zeros(2,dx);
```

```
i = 1;
j = 1;

do while i le dx;
   if d[.,i] == 0; di1[1,i] = ((0.1-dhat[i,.])/dbeta2[rd,.]);
elseif d[.,i] == 1; di1[1,i] = ((ln(2.50)-dhat[i,.])/dbeta2[rd,.]);
elseif d[.,i] == 2; di1[1,i] = ((ln(3.50)-dhat[i,.])/dbeta2[rd,.]);
          else; di1[1,i] = ((ln(5.00)-dhat[i,.])/dbeta2[rd,.]);
   endif;
   i = i+1;
endo;

do while j le dx;
   if d[.,j] == 0; di1[2,j] = ((ln(2.50)-dhat[j,.])/dbeta2[rd,.]);
elseif d[.,j] == 1; di1[2,j] = ((ln(3.50)-dhat[j,.])/dbeta2[rd,.]);
elseif d[.,j] == 2; di1[2,j] = ((ln(5.00)-dhat[j,.])/dbeta2[rd,.]);
          else; di1[2,j] = ((ln(15)-dhat[j,.])/dbeta2[rd,.]);
   endif;
   j = j+1;
endo;

proc integr_d(eps);
   local rd;
   rd = rows(dbeta2);
   retp((eps.*dbeta2[rd,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

dint = intquad1(&integr_d,di1);
_intord = 40;

dint1=(1./(1+exp(-di1)));
dint2=dint1[1,.]-dint1[2,.];

dint3=dint./dint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_d = dhat+dint3;

output file=f_dishes.out reset;

print lnrswg_d;

output off;
```

Appendix G
Estimation of Reservation Wages for Male Sample in Occupation Supermarket

```
/*        M_MARKET.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containing data.        */

open f1 = h:\thesis\data\data1\male\male.dat for read;
x = readr(f1,1441);

/*        Specifying data matrices, x1 contains independent variables, y1 – y4 contain
identifiers for the self-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the
reservation wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xs1 = x1~d2~d3~d4~c2~c3~c4~g2~g3~g4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y3,x1);
b3 = olsqr(y3,xs1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b3 = b3|0.5;

__title = "Logistic estimation of Supermarket for Males";

/*      Defining the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(s1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
s2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
s3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
s4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 1;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.          */

output file=sbeta_m.out reset;

{sbeta1,f1,g1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));

{sbeta2,f2,g2,h2,retc2} = maxprt(maxlik(xs1,0,&lli,b3));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{sbeta3,f3,g3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-squared =" prsq;

output off;

/*      Approximationg the error terms on the reservation wage equation.          */

sx = rows(xs1);
rs = rows(sbeta2);
vs = seqa(1,1,rs-1)';
sbeta4 = submat(sbeta2,vs,0);
shat = xs1*sbeta4;

s = (maxc(0*s1'|s2'|2*s3'|3*s4'))';
si1 = zeros(2,sx);
```

```
i = 1;
j = 1;

do while i le sx;
   if s[.,i] == 0; si1[2,i] = (((0.1)-shat[i,.])/sbeta2[rs,.]);
elseif s[.,i] == 1; si1[2,i] = ((ln(2.50)-shat[i,.])/sbeta2[rs,.]);
elseif s[.,i] == 2; si1[2,i] = ((ln(3.50)-shat[i,.])/sbeta2[rs,.]);
          else; si1[2,i] = ((ln(5.00)-shat[i,.])/sbeta2[rs,.]);
   endif;
   i = i+1;
endo;

do while j le sx;
   if s[.,j] == 0; si1[1,j] = ((ln(2.50)-shat[j,.])/sbeta2[rs,.]);
elseif s[.,j] == 1; si1[1,j] = ((ln(3.50)-shat[j,.])/sbeta2[rs,.]);
elseif s[.,j] == 2; si1[1,j] = ((ln(5.00)-shat[j,.])/sbeta2[rs,.]);
          else; si1[1,j] = ((ln(15)-shat[j,.])/sbeta2[rs,.]);
   endif;
   j = j+1;
endo;

proc integr_s(eps);
   local rs;
   rs = rows(sbeta2);
   retp((eps.*sbeta2[rs,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

sint = intquad1(&integr_s,si1);
_intord = 40;

sint1=(1./(1+exp(-si1)));
sint2=sint1[1,.]-sint1[2,.];

sint3=sint./sint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_s = shat+sint3;

output file=m_market.out reset;

print lnrswg_s;

output off;
```

Appendix H
Estimation of Reservation Wages for Female Sample in Occupation Supermarket

```
/*        F_MARKET.EST        */
new;
library gauss,user,maxlik;
#include maxlik.ext;
maxset;

/*        Opening file containg data.        */

open f1 = h:\thesis\data\data1\female\female.dat for read;
x = readr(f1,1535);

/*        Specifying data matrices, x1 contains independent variables, y1 – y4 contain
identifiers for the self-reported reservation wage category.        */

v0 = seqa(1,1,34)';
v1 = 35;
v2 = 36;
v3 = 37;
v4 = 38;

r = rows(x);
x1 = ones(r,1)~submat(x,0,v0);
y1 = submat(x,0,v1);
y2 = submat(x,0,v2);
y3 = submat(x,0,v3);
y4 = submat(x,0,v4);

/*        The following section specifies the dummy variables indicating the reservation
wage category for all four jobs.        */

d1 = zeros(r,1); d2 = zeros(r,1); d3 = zeros(r,1); d4 = zeros(r,1);
   d1 = (y1[.,1] .== zeros(r,1));
   d2 = (y1[.,1] .== ones(r,1));
   d3 = (y1[.,1] .== 2*ones(r,1));
   d4 = (y1[.,1] .== 3*ones(r,1));
c1 = zeros(r,1); c2 = zeros(r,1); c3 = zeros(r,1); c4 = zeros(r,1);
   c1 = (y2[.,1] .== zeros(r,1));
   c2 = (y2[.,1] .== ones(r,1));
   c3 = (y2[.,1] .== 2*ones(r,1));
   c4 = (y2[.,1] .== 3*ones(r,1));
s1 = zeros(r,1); s2 = zeros(r,1); s3 = zeros(r,1); s4 = zeros(r,1);
```

```
   s1 = (y3[.,1] .== zeros(r,1));
   s2 = (y3[.,1] .== ones(r,1));
   s3 = (y3[.,1] .== 2*ones(r,1));
   s4 = (y3[.,1] .== 3*ones(r,1));
g1 = zeros(r,1); g2 = zeros(r,1); g3 = zeros(r,1); g4 = zeros(r,1);
   g1 = (y4[.,1] .== zeros(r,1));
   g2 = (y4[.,1] .== ones(r,1));
   g3 = (y4[.,1] .== 2*ones(r,1));
   g4 = (y4[.,1] .== 3*ones(r,1));

xs1 = x1~d2~d3~d4~c2~c3~c4~g2~g3~g4;

/*      Creating starting points and other information for maximum likelihood
estimation.          */

b0 = olsqr(y3,x1);
b3 = olsqr(y3,xs1);

w1 = ln(2.50);
w2 = ln(3.50);
w3 = ln(5.00);
wc1 = w1*ones(r,1);
wc2 = w2*ones(r,1);
wc3 = w3*ones(r,1);

b0 = b0|0.5;
b3 = b3|0.5;

__title = "Logistic estimation of Supermarket for Females";

/*      Defininf the log-likelihood function.          */

proc lli(b,x);
   local t,u,v1,beta;
   t = rows(x);
   u = rows(b);
   v1 = seqa(1,1,rows(b)-1);
   beta = submat(b,v1,0);
   fn logist(m) = 1/(1 + exp(-m));
   retp(s1'*ln(maxc((logist((wc1-x*beta)/b[u,.]))'|1e-10*ones(1,t)))
+
s2'*ln(maxc((logist((wc2-x*beta)/b[u,.])-logist((wc1-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
```

```
s3'*ln(maxc((logist((wc3-x*beta)/b[u,.])-logist((wc2-x*beta)/b[u,.]))'
|1e-10*ones(1,t)))
+
s4'*ln(maxc((ones(t,1)-logist((wc3-x*beta)/b[u,.]))'|1e-10*ones(1,t))));
endp;

/*      Specifying the settings for the optimization process.          */

_mlstmth = "steep one nohess";
_mlmdmth = "bfgs stepbt hess";
_mldfct = 0.01;
_mlditer = 50;
_mlcovp = 0;
_mlgtol = 0.001;
_mlmiter = 500;

/*      Estimating reservation wages, equations 3 and 10 in Chapter 4 and generating
output.          */

output file=sbeta_f.out on;

{sbeta1,f1,g1,h1,retc1} = maxprt(maxlik(x1,0,&lli,b0));
/*
{sbeta2,f2,g2,h2,retc2} = maxprt(maxlik(xs1,0,&lli,b3));

xa1 = ones(r,1);
beta2 = 0.5*ones(2,1);
{sbeta3,f3,g3,h3,retc3} = maxprt(maxlik(xa1,0,&lli,beta2));

prsq = 1 - (f2/f3);
print "Psuedo R-square =" prsq;

output off;

/*      Approximating the error terms on the reservation wage equation.          */

sx = rows(xs1);
rs = rows(sbeta2);
vs = seqa(1,1,rs-1)';
sbeta4 = submat(sbeta2,vs,0);
shat = xs1*sbeta4;

s = (maxc(0*s1'|s2'|2*s3'|3*s4'))';
si1 = zeros(2,sx);
```

```
i = 1;
j = 1;

do while i le sx;
   if s[.,i] == 0; si1[2,i] = (((0.1)-shat[i,.])/sbeta2[rs,.]);
elseif s[.,i] == 1; si1[2,i] = ((ln(2.50)-shat[i,.])/sbeta2[rs,.]);
elseif s[.,i] == 2; si1[2,i] = ((ln(3.50)-shat[i,.])/sbeta2[rs,.]);
          else; si1[2,i] = ((ln(5.00)-shat[i,.])/sbeta2[rs,.]);
   endif;
   i = i+1;
endo;

do while j le sx;
   if s[.,j] == 0; si1[1,j] = ((ln(2.50)-shat[j,.])/sbeta2[rs,.]);
elseif s[.,j] == 1; si1[1,j] = ((ln(3.50)-shat[j,.])/sbeta2[rs,.]);
elseif s[.,j] == 2; si1[1,j] = ((ln(5.00)-shat[j,.])/sbeta2[rs,.]);
          else; si1[1,j] = ((ln(15)-shat[j,.])/sbeta2[rs,.]);
   endif;
   j = j+1;
endo;

proc integr_s(eps);
   local rs;
   rs = rows(sbeta2);
   retp((eps.*sbeta2[rs,.]).*(exp(eps))./((1+exp(eps))^2));
endp;

sint = intquad1(&integr_s,si1);
_intord = 40;

sint1=(1./(1+exp(-si1)));
sint2=sint1[1..]-sint1[2,.];

sint3=sint./sint2';

/*      Creating final estimate of log-reservation wage and generating output.      */

lnrswg_s = shat+sint3;

output file=f_market.out reset;

print lnrswg_s;

output off;*/
```

Appendix I
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
INTERCEPT	0.526*	0.5235*	0.9205*	-0.0593
	(2.558)	(2.02)	(3.587)	(0.258)
AGE	0.0808	0.0085	0.4951*	-0.1991
	(0.434)	(0.023)	(2.392)	(0.916)
GRADE	-0.094	0.0343	-0.0368	0.9892*
	(0.523)	(0.124)	(0.091)	(4.603)
HSGRAD	0.0547	-0.0209	0.014	-0.0986*
	(1.494)	(0.494)	(0.219)	(2.258)
HSVOCCOM	-0.0182	-0.0005	0.0178	-0.004
	(0.627)	(0.013)	(0.488)	(0.115)
HSGENPRG	-0.0058	-0.0311	0.0396	0.0156
	(0.252)	(1.183)	(1.393)	(0.565)
PCTWRK78	0.0186	-0.0329	0.008	0.0566
	(0.637)	(0.986)	(0.241)	(1.629)
PCTWRK80	0.0707*	0.0036	0.0006	-0.0155
	(2.636)	(0.121)	(0.015)	(0.477)
MSTAT	-0.041	0.0374	0.005	0.0754
	(0.882)	(0.561)	(0.105)	(1.32)
BLACK	-0.0262	0.0058	-0.0759*	-0.0546*
	(1.099)	(0.217)	(2.884)	(1.966)

Appendix I (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
HISPANIC	0.0091	0.0325	-0.0716*	0.0312
	(0.312)	(1.04)	(2.356)	(0.957)
MOMSEDHS	-0.0382	0.016	0.0029	0.0581*
	(1.758)	(0.581)	(0.094)	(2.22)
MOMSEDNA	-0.105*	0.0706	-0.0669	0.1257*
	(2.568)	(1.463)	(1.461)	(2.576)
DADSEDHS	0.0499*	-0.0013	-0.0579*	-0.0047
	(2.225)	(0.043)	(2.188)	(0.176)
DADSEDNA	0.0212	-0.0114	-0.0258	0.0055
	(0.747)	(0.302)	(0.803)	(0.156)
WELFARE	-0.0192	0.0381	-0.0309	-0.0088
	(0.732)	(1.262)	(1.026)	(0.292)
ACPTWELF	0.0189	0.0007	-0.0238	0.011
	(0.985)	(0.03)	(1.033)	(0.493)
ROTTER	0.0218	-0.015	0.0476	0.0271
	(0.622)	(0.385)	(1.235)	(0.661)
KNOWWORK	-0.066	0.0896	0.0064	0.0873
	(1.496)	(1.815)	(0.128)	(1.666)
EDUCGOAL	0.1038	0.1837	-0.1563	-0.1905
	(1.157)	(1.74)	(1.344)	(1.805)

Appendix I (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
NOTENRL	-0.0088	0.0242	-0.0207	-0.036
	(0.283)	(0.614)	(0.574)	(0.983)
EMPLOYED	-0.0049	-0.0372	0.0425	-0.0789
	(0.105)	(0.691)	(0.781)	(1.432)
OLF	-0.0232	0.0261	0.0606	0.0194
	(0.878)	(0.857)	(1.905)	(0.594)
LOGWAGES	0.0414	0.0438	0.0002	0.0901*
	(1.265)	(1.162)	(0.004)	(2.383)
WAGENA	0.0488	-0.0015	0.0282	0.0652
	(0.879)	(0.022)	(0.44)	(1.019)
WELFNA	0.1117*	-0.1044	-0.0678	0.0463
	(2.026)	(1.808)	(1.117)	(0.742)
HSTYPENA	-0.0644	0.0452	0.0282	0.0495
	(1.422)	(0.848)	(0.57)	(0.901)
ROTTERNA	0.0544	-0.111	0.1348	0.1865
	(0.638)	(1.182)	(1.22)	(1.646)
URATE	-0.1097	0.2388	-0.2197	0.0553
	(0.974)	(1.545)	(1.757)	(0.462)
URBAN	0.0196	0.0324	-0.0035	-0.0514
	(0.274)	(0.379)	(0.031)	(0.606)

Appendix I (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
POORCNTY	1.2197*	-1.2158*	-1.2205*	1.0433*
	(3.191)	(2.797)	(2.882)	(2.378)
PCAPINC	0.0353	-0.1161	-0.6196*	0.456*
	(0.186)	(0.519)	(2.804)	(2.068)
NCENTRAL	0.0314	0.0026	-0.0231	0.0295
	(1.178)	(0.061)	(0.754)	(1.01)
SOUTH	-0.0463	0.0664	-0.0494	0.0126
	(1.297)	(1.236)	(1.153)	(0.334)
WEST	0.0092	0.004	-0.06	0.0679
	(0.286)	(0.077)	(1.747)	(1.921)
D.D2	---	0.28*	0.0337	0.133*
		(8.355)	(0.856)	(3.426)
D.D3	---	0.5263*	0.0664	0.2273*
		(14.346)	(1.514)	(5.332)
D.D4	---	0.8375*	0.2024*	0.4474*
		(19.76)	(4.208)	(9.208)
D.C2	0.2554*	---	0.1135*	0.0833*
	(8.219)		(2.797)	(2.03)
D.C3	0.3977*	---	0.1488*	0.1251*
	(12.553)		(3.496)	(2.895)

Appendix I (continued)
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
D.C4	0.6733*	---	0.2689*	0.1712*
	(19.63)		(5.91)	(3.634)
D.S2	0.0198	0.0795*	---	0.2462*
	(0.722)	(2.532)		(8.023)
D.S3	0.0832*	0.0843*	---	0.3819*
	(2.693)	(2.338)		(10.823)
D.S4	0.1996*	0.2784*	---	0.6765*
	(5.366)	(6.344)		(15.582)
D.B2	0.0639*	0.0669	0.2741*	---
	(2.183)	(1.952)	(8.035)	
D.B3	0.167*	0.1009*	0.3957*	---
	(5.168)	(2.625)	(10.627)	
D.B4	0.3194*	0.1455*	0.6530	---
	(8.821)	(3.432)	(16.260)	
SCALE(σ_j)	0.1459*	0.1704*	0.1727*	0.1766*
	(27.248)	(26.815)	(27.568)	(26.759)
Pseudo R^2	0.4198	0.3574	0.3411	0.3783
Observations	1441	1441	1441	1441

Appendix I
Estimates of the Determinants of Males Self-Reported Reservation Wages, Equation (10) in Chapter 4.

Absolute Value of Asymptotic t-statistics in parentheses.

Notes: * Indicates the parameter estimate is significantly different from zero (two-tailed test) at the 5% level

Pseudo R^2 is calculated as $1-[(\log L_\Omega)/(\log L_\omega)]$, where L_Ω is the maximized value of the unrestricted likelihood function and L_ω is the maximized value of the likelihood function with no explanatory variables (Maddala, 1983, p. 40).

"D.X#" identifies the dummy variable which takes the value 1 if the respondent's reservation wage for job "X" is reported to be in the "#" wage category. "X" corresponds to the first letter of the job category.

Appendix J
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
INTERCEPT	0.6539*	0.5424*	0.2739	-0.5123*
	(2.493)	(2.045)	(1.135)	(2.408)
AGE	-0.2397	-0.2459	0.2563	0.862*
	(0.979)	(1.004)	(1.28)	(4.51)
GRADE	0.1025	0.114	0.0089	0.2086*
	(0.413)	(0.471)	(0.054)	(2.393)
HSGRAD	0.1326*	-0.0591	0.12*	-0.0796*
	(2.775)	(1.217)	(2.783)	(2.549)
HSVOCCOM	-0.0374	0.0505	0.011	0.0149
	(1.043)	(1.357)	(0.305)	(0.558)
HSGENPRG	-0.0429	-0.0213	0.0245	0.016
	(1.506)	(0.728)	(0.865)	(0.632)
PCTWRK78	-0.0286	0.0643	-0.0357	0.1073*
	(0.737)	(1.612)	(1.042)	(3.261)
PCTWRK80	0.0528	-0.0068	0.0253	0.0201
	(1.586)	(0.194)	(0.814)	(0.728)
MSTAT	0.0456	-0.0488	0.0221	0.057*
	(1.151)	(1.261)	(0.664)	(2.152)
BLACK	-0.0666*	0.0441	-0.0154	-0.0498*
	(2.172)	(1.425)	(0.538)	(2.001)

Appendix J (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
HISPANIC	0.0219	0.1109*	-0.0119	-0.0331
	(0.61)	(2.997)	(0.363)	(1.238)
MOMSEDHS	-0.0025	-0.008	-0.0001	0.0144
	(0.093)	(0.287)	(0.003)	(0.605)
MOMSEDNA	-0.0204	0.0595	-0.1193*	0.0515
	(0.42)	(1.15)	(2.469)	(1.451)
DADSEDHS	0.001	-0.0065	-0.0405	0.0421
	(0.034)	(0.215)	(1.562)	(1.726)
DADSEDNA	0.0297	0.014	-0.011	-0.026
	(0.84)	(0.397)	(0.322)	(0.986)
WELFARE	0.0052	-0.0326	-0.0297	0.0109
	(0.155)	(0.967)	(0.943)	(0.42)
ACPTWELF	0.0127	0.023	-0.0138	0.0358
	(0.524)	(0.934)	(0.616)	(1.438)
ROTTER	0.0007	-0.0056	-0.046	0.1054*
	(0.016)	(0.119)	(1.154)	(3.05)
KNOWWORK	0.1254*	-0.0596	-0.0406	0.0455
	(2.194)	(0.993)	(0.793)	(1.571)
EDUCGOAL	0.0796	0.161	0.05	-0.1193*
	(0.684)	(1.348)	(0.467)	(3.491)

Appendix J (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
NOTENRL	0.0592	-0.04	0.0446	-0.0227
	(1.431)	(0.944)	(1.131)	(0.761)
EMPLOYED	0.0765	-0.1157*	0.0556	0.0361
	(1.253)	(2.246)	(1.202)	(1.126)
OLF	-0.039	0.0246	0.065*	0.0788*
	(1.162)	(0.714)	(1.98)	(2.847)
LOGWAGES	-0.0456	0.0947*	-0.005	-0.0092
	(1.007)	(2.626)	(0.163)	(0.373)
WAGENA	-0.0192	0.1545*	0.022	-0.0078
	(0.221)	(2.166)	(0.365)	(0.254)
WELFNA	-0.0754	0.1245	0.0577	0.0157
	(0.835)	(1.311)	(0.727)	(0.357)
HSTYPENA	0.0138	-0.0736	0.0801	0.0092
	(0.234)	(1.201)	(1.363)	(0.283)
ROTTERNA	0.253	0.0776	0.032	0.1285*
	(1.341)	(0.411)	(0.153)	(2.01)
URATE	-0.2341	0.4323*	0.0349	0.0187
	(1.627)	(2.908)	(0.251)	(0.157)
URBAN	0.005	-0.0713	0.1209	0.0069
	(0.04)	(0.776)	(1.493)	(0.121)

Appendix J (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
POORCNTY	0.9451*	0.1569	-0.0884	-0.2605*
	(2.005)	(0.334)	(0.219)	(3.008)
PCAPINC	0.1876	0.2823	-0.1054	0.0286
	(0.755)	(1.181)	(0.499)	(0.363)
NCENTRAL	-0.0357	0.0182	-0.0374	0.1*
	(1.062)	(0.531)	(1.219)	(3.397)
SOUTH	-0.0618	0.0462	-0.0644	0.1282*
	(1.411)	(1.012)	(1.605)	(3.675)
WEST	-0.0678	0.0327	-0.0786*	0.1739*
	(1.749)	(0.812)	(2.242)	(5.505)
D.D2	---	0.2876*	0.1089*	0.0595
		(7.435)	(2.399)	(1.738)
D.D3	---	0.4845*	0.0769	0.1706*
		(12.164)	(1.588)	(4.862)
D.D4	---	0.7761*	0.1637*	0.3804*
		(17.989)	(3.298)	(10.835)
D.C2	0.2808*	---	-0.0473	0.2439*
	(6.808)		(0.843)	(6.694)
D.C3	0.3895*	---	0.0099	0.2864*
	(9.582)		(0.156)	(8.073)

Appendix J (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

Variable	Job Type			
	Dishes	Clean	Supermarket	Burgers
(1)	(2)	(3)	(4)	(5)
D.C4	0.7358*	---	0.0536	0.3928*
	(17.313)		(0.9)	(11.441)
D.S2	0.0278	-0.0177	---	0.2701*
	(0.898)	(0.533)		(8.962)
D.S3	0.0716	0.045	---	0.389*
	(1.934)	(1.161)		(10.947)
D.S4	0.2227*	0.1696*	---	0.8855*
	(4.364)	(3.242)		(16.756)
D.B2	0.100*7	0.1352*	0.3678*	---
	(2.8)	(3.721)	(9.709)	
D.B3	0.234*	0.1883*	0.4329*	---
	(5.8)	(4.522)	(10.16)	
D.B4	0.3653*	0.3509*	0.7042*	---
	(8.653)	(7.873)	(16.865)	
SCALE(σ_J)	0.1981*	0.205*	0.1958*	0.2123*
	(27.316)	(27.241)	(31.175)	(25.599)
Pseudo R^2	0.3167	0.2842	0.2206	0.3224
Observations	1535	1535	1535	1535

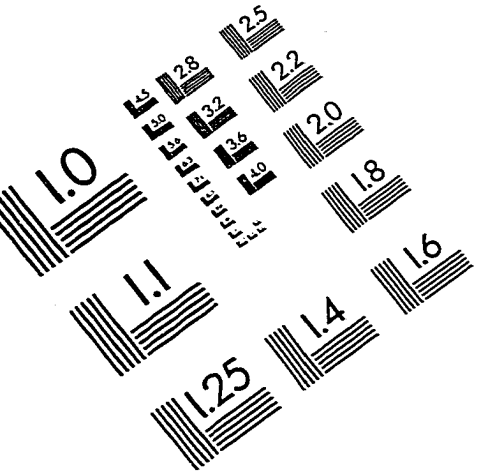
Appendix J (continued)
Estimates of the Determinants of Females Self-Reported Reservation Wages, Equation (10) in Chapter 4.
Absolute Value of Asymptotic t-statistics in parentheses.

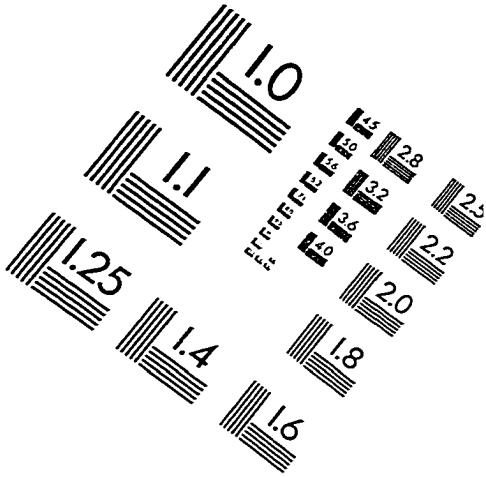
Notes: * Indicates the parameter estimate is significantly different from zero (two-tailed test) at the 5% level

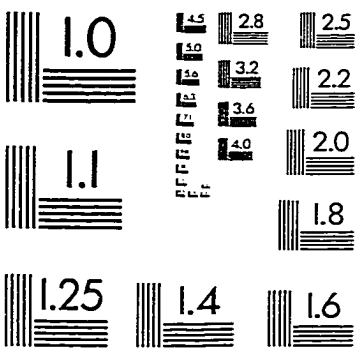
Pseudo R^2 is calculated as 1-[(log L_Ω)/(log L_ω)], where L_Ω is the maximized value of the unrestricted likelihood function and L_ω is the maximized value of the likelihood function with no explanatory variables (Maddala, 1983, p. 40).

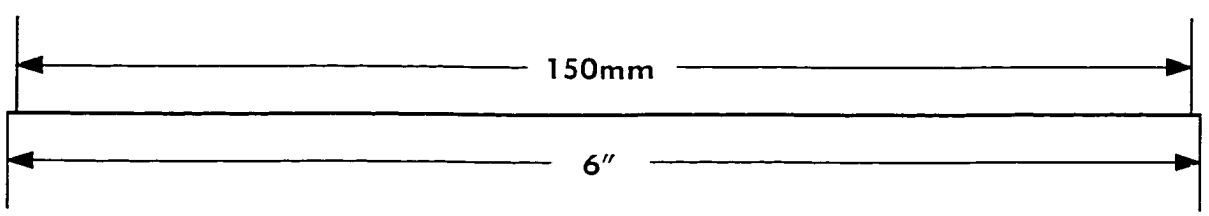
"D.X#" identifies the dummy variable which takes the value 1 if the respondent's reservation wage for job "X" is reported to be in the "#" wage category. "X" corresponds to the first letter of the job category.

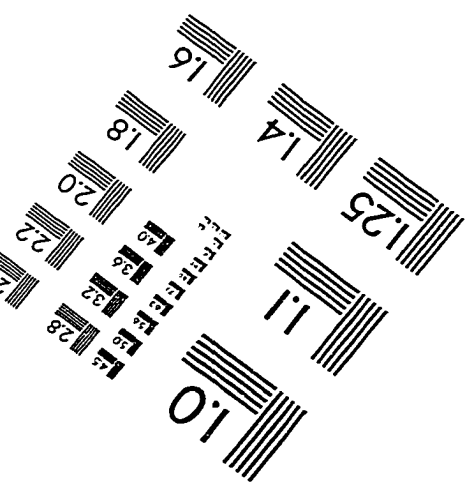
IMAGE EVALUATION TEST TARGET (QA-3)

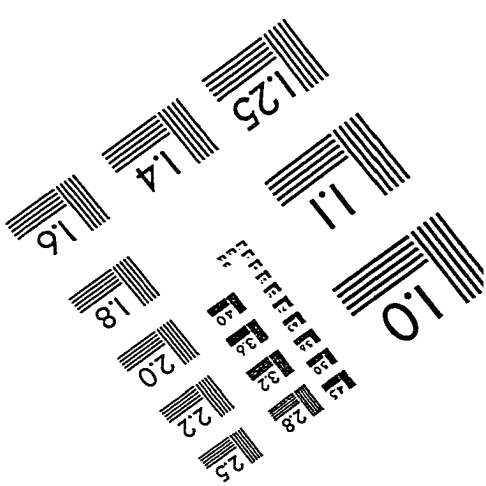












APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989