THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION

OFFER ARRIVALS VERSUS ACCEPTANCE: INTERPRETING DEMOGRAPHIC REEMPLOYMENT PATTERNS IN THE SEARCH FRAMEWORK

No. 104

T. J. Devine
Pennsylvania State University

U. S. Department of Commerce BUREAU OF THE CENSUS

ACKNOWLEDGEMENTS

I would like to thank Nick Kiefer. George Jakubson, and David Easley for many helpful discussions and comments on this work. Comments from Ken Burdett, Lars Muus, and members of the Cornell Labor Workshop were also helpful. Most of the computing was done at the Cornell national Supercomputer Facility, funded in part by the National Science Foundation. The Center for Analytic Economics at Cornell is also thanked for its support.

This Working Paper also appears in the Working Paper Series (No. 4-88-2) of the Department of Economics, the Pennsylvania State University. It was originally completed in April 1988 and revised in October 1989. The author has kindly permitted the paper to be made available to the SIPP user community through the SIPP Working Paper Series. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau.

TABLE OF CONTENTS

TABLES

FIGURES

The structure of the stationary job search model is now familiar. The unemployed worker searches each period for wage offers from some distribution of possibilities; with some probability, an offer is secured. The worker knows both the distribution of offers and the probability of receiving an offer. Uncertainty exists because the worker does not know which firms are making which offers. When an offer is received, the worker must choose between accepting the offer or continuing the search. Suppose that the worker's objective is maximization of the discounted value of lifetime income, that unemployment income net of search costs does not change as a spell continues, and that the worker expects to hold any job taken for a long time. The optimal policy for the worker in this case is to accept the first offer that exceeds a time-invariant reservation wage--this being the wage which equates the marginal cost and marginal benefit of continued search.

Given the model's simplicity, it seems unreasonable to expect it to provide a precise description of an individual worker's experience. Complicating the model--by adding structure to the worker's environment or behavior--would thus seem an appropriate strategy when attempting to explain observed variation across workers in the transition rate into employment. Suppose, however, that the objective is to explain variation across major groups that comprise the labor force. One might maintain the simple structure of the model at the individual worker level, but allow for the existence of only a few worker types, i.e., stop somewhere between the macro and micro levels of aggregation and see if the simple model is useful for explaining observed variation across major groups in the labor force. This is the approach taken here.

The basic job search model implies that unemployment durations are determined by a constant transition rate--the rate at which acceptable offers are received. This factors into two components: the rate at which workers receive offers--the *arrival rate*--and the probability that workers accept offers once received--the *acceptance probability*. My objective is to distinguish empirically the relative roles of these two factors in producing observed variation in the transition rate into employment across different worker types. Specifically, I investigate variation among groups that differ by age, race, and sex.[2]

Others have studied variation in transition rates across individuals in terms of arrival rates and acceptance probabilities. Toward this end, parametric functions of individual characteristics and local labor market conditions have been specified; parameter restrictions have then been imposed to identify the arrival rate and acceptance probability upon estimation. In some studies, unusual data on numbers of offers and reservation wages are also relied upon for identification. In general, the results of such attempts are not encouraging--poor fits are obtained with the individual level data. Since my interest is in explaining variation across groups of workers, I carry out estimation using grouped duration and accepted wage data--treating workers within each subsample as a random sample from a homogeneous population. This approach allows me to estimate the transition rate, arrival rate, and acceptance probability for each group using methods that require replication (but do not require unusual data on reservation wages or actual frequency of offer receipt). Relationships between the key variables are then studied using the results for all groups together. In this sense, my partition of the sample can be viewed as the identifying restriction required to determine the relative roles of the two factors in producing variation in transition rates across workers.[3]

Section 1 sets out a search model that incorporates the existence of workers of different types. Section 2 presents the grouped data estimation approach. Results from application of the approach to recently released micro data for a large sample of U.S. workers are presented in Section 4. The sample analyzed here is taken from the 1984 Panel of the Survey of Income and Program Participation (SIPP). Given that the longitudinal features of SIPP have not been exploited previously, an overview of the SIPP is provided along with a description of the data set used here in Section 3. Section 5 concludes.

The results presented here provide evidence of some variation in the acceptance probability across groups, but there is evidence of substantial variation in the arrival rate across groups. Substantial variation in the transition rate across groups thus appears to be directly related to arrival rate variation. This result appears robust with respect to both errors in measurement and alternative parametric assumptions for the distributions of wage offers.

Section 1. The Model

The labor force is viewed here as consisting of distinct groups of workers. Within each group, workers are assumed homogeneous. Whether this assumption is reasonable depends on the definition of groups. Obviously, this assumption holds when groups are individuals, perhaps not when the group is the total labor force. The idea here is to locate a reasonable middle ground. At the individual level within a group, I adhere to the standard theoretical job search framework. An individual's labor market history is therefore modeled as a stochastic process which moves among labor market states in response to random events--job offers and layoffs for unemployed and employed workers, respectively.[4]

Precisely, letting groups be indexed by c, the individual worker seeks to maximize expected lifetime income, discounted to the present over an infinite horizon at some constant positive rate r_c. When unemployed, the worker searches for job offers and offers arrive according to a time-homogeneous Poisson process with parameter δ_c, referred to as the arrival rate.[5] A job offer is summarized by a wage rate w that will be received continuously over tenure of employment, if accepted, and successive job offers are independent realizations from a known wage offer distribution with finite mean μ_c, distribution function $F_c(w)$, and density $f_c(w)$. Once employed, a worker may be laid off; the occurrence of these layoffs follows a Poisson process with parameter a_c. Finally, the income flow while unemployed (net of any search costs) is fixed over the course of a given spell at rate b_c and there is no on-the-job search. Under these assumptions, the optimal acceptance/rejection strategy for the worker is a time-invariant reservation wage policy: accept $w \geq w_c^r$, where the reservation wage wc^r is defined by equating the expected present value of employment and the expected present value of continued search.

My interest is in the empirical implications of this simple model. Let τ_c denote the instantaneous probability that an individual of type c will become reemployed. This is simply the instantaneous probability that an acceptable offer will be received by this individual,

$$\tau_c = \delta_c \pi_c(w_c^r) \ .$$

The first term is the arrival rate. The second term is the conditional probability that an offer, once received, will be accepted under the worker's optimal policy,

$$\pi_c(w_c^r) = \int_{w_c^r}^{\infty} dF_c(w)$$

$$= 1 - F_c(w_c^r) \ .$$

This is the acceptance probability. The transition rate between the states of unemployment and employment, τ_c (also referred to as the instantaneous reemployment probability or hazard rate) does not depend on elapsed duration, nor does it depend on calendar time--because neither the preferences of the worker, nor the environment depends on these measures of time. This in turn has implications for the distribution of unemployment spell durations T_c. Completed durations have an exponential distribution with parameter τ_c and mean $1/\tau_c$.[6]

In looking across groups of workers of different types, variation in mean unemployment spell lengths may thus be discussed in terms of variation in transition rates τ_c. In turn, variation in transition rates across types may be attributed to variation in arrival rates δ_c or acceptance probabilities $\pi_c(w_c^r)$. The model implies nothing about the relative roles of these two variables, however. The model instead renders this an empirical question.

While fairly standard in structural empirical job search studies, the assumptions invoked above are restrictive.[7] The extent to which each may be relaxed without greatly affecting the basic empirical implications of the model is not far. In particular, with any source of time variation in τ_c--such as variation in b_c or δ_c--we lose the exponential distribution for durations. This may explain, at least in part, the limited results from structural studies that attempt to fit stationary models to individual level data. The intention here is to see if the model is nevertheless useful as a description of average experience among individuals within groups--defined at a level between the micro and macro levels--as opposed to serving as a precise description of each group member's experience.[8]

Section 2. The Empirical Approach

Only a relative few empirical studies -- Mortensen and Neumann (1984), Narendranathan and Nickell (1985), Ridder and Gorter (1986), and Blau and Robbins (1986) -- have focused on distinguishing arrival rates and acceptance probabilities. These studies all focus on experience at the individual worker level. Accordingly, their econometric approaches involve parameterization of each structural element in the model. That is, functions defined over individual characteristics, income variables, and labor market conditions are specified for the arrival rate, parameters of the offer distribution, etc. A number of somewhat arbitrary restrictions are then suggested for identification. In some cases, unusual data are also relied upon for the purpose of identification (numbers of offers in the study by Blau and Robins (1986) and reported reservation wages and minimum wage offers in the study by Ridder and Gorter (1986)). Generally, consistency checks on such data are not favorable to the interpretation used in estimation, which calls their use into question.

My approach to the data instead involves working directly with a partition of a sample from the labor force based on demographic characteristics. Specifically, for each group c within a partition of the labor force, there exists a set of basic search parameters $\{\tau_c, \delta_c, F_c(w), w_c^r\}$. I estimate these parameters using accepted wage and duration data for the sample of workers of type c, alone, and then interpret the results as representative for all individuals having the characteristics that define the group. Beyond specification of a partition, the only parameterization required at the empirical stage is specification of a parametric family of wage offer distributions $\{F(w|\theta_c, \theta_c \in \Theta\}$. This is unavoidable in the absence of rejected offer data (which is the typical situation and the particular situation faced here).

With estimates of the arrival rate and acceptance probability for each group, I can consider whether observed variation in transition rates across groups reflects systematic variation in arrival rates or acceptance probabilities or both. If the data support such relationships, it is important to know whether this finding can be attributed to the particular family of offer distributions specified. Experimentation with a variety of distributional assumptions consequently seems an appropriate empirical strategy. Specification diagnostics are then used to check the sensitivity of the results for $\pi_c(w^r{}_c)$ and δ_c, and questions concerning their relative roles are addressed on the basis of the overall results. The precise strategy I follow to estimate the elements of the vector $\{\tau_c, \delta_c, \{F(w|\theta_c, \theta_c \in \Theta\}, w_c{}^r\}$ for each group c is the following:

The Reservation Wage

Assuming that the wages observed for individuals in group c are realizations of independently and identically distributed random variables with distribution function $F_c(w)$, a number of consistent estimators $w_c{}^{r*}$ for $w_c{}^r$ based on accepted wages are available. In particular, any of the first m order statistics, m fixed, and their averages represent strongly consistent estimators. I use the average of the first two order statistics.[9]

The Transition Rate

The search model set out above implies that completed unemployment spell lengths of all workers in group c are independently and identically distributed according to an exponential distribution with parameter τ_c. The maximum likelihood estimate of the transition rate for group c therefore involves a straightforward calculation using data on spell durations for the group c sample. Let $d_{ci} = 1$ if observation i in the group c sample is censored and $d_{ci} = 0$ otherwise. Let N_c denote the group c sample size. Then the maximum likelihood estimator is given by

$$\tau_c^* = \frac{\sum_{i=1}^{N_c} d_{ci}}{\sum_{i=1}^{N_c} t_{ci}} .$$

The Offer Arrival Rate

With estimates of the reservation wage and the transition rate for group c, if I know the distribution $F_c(w)$, then I need only assume that the arrival rate and offer distribution are stochastically independent to calculate a consistent estimate of the arrival rate. That is, since $\tau_c = \delta_c \pi_c(w_c^r)$, I can use

$$\delta_c^* = \tau_c^* / \pi_c(w_c^{r*}),$$

as an estimator. I do not know $F_c(w)$ or, equivalently $\pi_c(w) = 1 - F_c(w)$, but must estimate this instead. I turn to this next.

The Offer Distribution

Specification of a parametric family of wage offer distributions is unavoidable for two reasons if identification of the arrival rate δ_c is desired. First, economic theory says very little about the true offer distribution. Second, the true offer distribution $F_c(w)$ faced by the type c workers cannot be determined uniquely from my sample wage data using nonparametric methods alone -- regardless of the group c sample size--because only accepted wages are observed. That is, observed wages are drawn from distributions truncated at the reservation wage w_c^r with density

$$f_c(w|w \geq w_c^r) = \frac{f_c(w)}{\pi_c(w_c^r)}, \quad w \geq w_c^r$$

$$= 0 \quad , \text{otherwise} .$$

Use of nonparametric methods requires data on $w < w_c^r$, i.e., information regarding the mass below the reservation wage.[10]

Given this situation, I assume that $F_c(w)$ is a member of a parametric family $\{F(w|\theta_c), \theta_c \in \Theta\}$ and use the observations from the truncated distribution to estimate the parameters θ_c. The set from which this parametric family may be chosen is not without restrictions. Obviously, the vector of parameters must be estimable from the data on the accepted wages, i.e. no element can depend on anything below w_c^r.[11] The normal and gamma families represent candidate families and both are attractive in that they are two parameter families and allow for different shapes. Since my estimator for the arrival rate δ_c^* may be sensitive to choice of family, I estimate of the full set of parameters under both specifications.

The parameters θ_c are estimated using the method of moments. Precisely, theoretical moments $m(\theta_c)$ and sample moments S_c for the truncated distribution are equated, yielding a system of k nonlinear simultaneous equations in θ_c

$$S_c = m(\theta_c) \ ,$$

where the choice of k satisfies $k \geq p$, the dimension of θ_c. Consistent estimates are then obtained by solving the minimum distance problem

$$\min_{\substack{\theta \\ \theta \in \Theta}} \ D(\theta) = [S_c - m(\theta)]' \ A \ [S_c - m(\theta)] \ ,$$

where A is a consistent estimate for the inverse of the asymptotic covariance matrix for S_c.[12] By working with more than two moments of the accepted offer distribution, the system of equations is overidentified. A specification check--in the form of a test of the overidentifying restriction--is therefore available under each distributional assumption.

As with any identification method, there is a price involved with taking a grouped data approach. I cannot infer the effects of heterogeneity remaining within each group at each level of aggregation. Marginal effects of particular variables on turnover (such as the effect of a single year increase in age) are not ascertained. However, results obtained by others suggest that pushing the stationary job search model to explain behavior at the individual level may be pushing it too far. My objective here is to determine whether the simple model is useful for explaining observed variation across major groups of workers in the labor force, as opposed to variation across individual workers. The grouped data method delivers information on just that.

Section 3. Data

The data are taken from the Public Use Files for the 1984 Panel of the Survey of Income and Program Participation (SIPP), a nationwide longitudinal survey conducted by the Bureau of the Census. The SIPP is an attractive data source for empirical labor economics for several reasons. Labor force activity and income data are available in finer detail than that offered by alternative sources for the U.S. The data are also collected more frequently (every four months, as opposed to annually). The Panel (i.e., sample, using the SIPP terminology) is relatively large and it is also representative for the U.S. The less desirable feature of the SIPP data is the relatively complex structure of the files available to the public. For example, labor force activity information is reported week by week within Waves and the data are available Wave by Wave (i.e., for each four month survey period). Merging data for individuals across Waves is less than straightforward, but necessary to exploit SIPP's longitudinal features.[13]

I work with data from the first four Waves of the survey (i.e., sixteen months) for a sample of 5,214 workers. Specifically, the sample consists of those of all individuals who: (i) experienced an initialized spell of joblessness (new entrants are thus excluded), (ii) did not report having a job in either an agricultural occupation or an agricultural industry, (iii) either worked full-time hours or reported part-time hours were due to economic reasons when employed, (iv) remained age 64 or less at the end of the first completed spell of joblessness or the end of the sixteen month period considered, and (v) were neither disabled nor self-employed during the survey period. Table 3.1 presents a summary of the sample.[14]

The two key variables in my analysis are accepted wages and unemployment spell lengths.[15] Neither is reported directly in the SIPP, but numerous related items are reported. A number of decisions were consequently required on my part. Durations are measured as weeks of "joblessness," where being "with a job" is defined as having an arrangement with an employer for regular work.[16] The SIPP provides answers to questions pertaining to hourly wages for hourly workers and, where available, these data are exploited. Average hourly earnings, based on accepted monthly earnings and weekly hours, are used in remaining cases.[17]

Section 4. Empirical Specification and Results

The results presented here are for a partition of the labor force by race, sex, and age. Using an age partition of 16-19, 20-24, 25-44, 45-64 years, there are sixteen groups (4 x 2 x 2). (For reference, the complete classification scheme is given in Table 4.1). Table 4.2 presents estimates for the transition rate between unemployment and employment τ_c (TAU) for each group and the expected jobless spell lengths based on these estimates (E(DUR)). Note that these estimates do not depend on a distributional assumption for offers, but do rely on the exponential specification for durations. The number of duration observations (N DUR) and number of wage observations (N WAGE) are reported, since these vary across groups. (Obviously more confidence can be placed in results for white workers and younger workers.)

Some very clear age and race patterns appear in the estimates for the group transition rates. First, older workers become reemployed much more slowly than younger workers. Nonwhite workers also leave more slowly than their white counterparts, although the contrast is less sharp for female teens (16-19) and young adult males (20-24). As for gender differences, white males within each age group tend to move out more quickly than their female counterparts, but the difference is substantial only for prime-age workers (25-44). Among nonwhite workers, on the other hand, this gender difference appears only for the young adult and prime-age nonwhite groups and, even for these age groups, the nonwhite gender differences are relatively small.

Overall, the demographic pattern exhibited in the transition rate estimates is roughly consistent with those based on CPS gross flow data (e.g., Ehrenberg (1981)). My interest is in determining which of the two potential factors -- the arrival rate or the acceptance probability -- plays the greater role in producing this pattern. Tables 4.3 and 4.4 present the results for the search parameters $\{\tau_c, \delta_c, \pi_c(w_c^r)\}$ from estimation under the gamma and normal offer specifications. The average of the first two order statistics (WR2A) serves as the reservation wage estimator in both cases. Under both specifications, the arrival rate appears to be the dominant factor. This can be gathered from careful inspection of the estimates for the offer probabilities (NPI and GPI for the normal and gamma, respectively) and the arrival rates (NDELTA and GDELTA for the normal and gamma, respectively) or the plots of the transition rate τ_c estimates against the arrival rate δ_c estimates and against the acceptance probabilities $\pi_c(w_c^r)$ (Figures 4.1-2 and 4.3-4 for the gamma and normal offer specifications, respectively).[18]

The positive, essentially linear relationship exhibited in the plots of TAU against the arrival rates GDELTA and NDELTA strongly suggests that groups with higher transition rates are groups with higher arrival rates. The simple correlations between the τ_c and δ_c estimates are 0.998 under the gamma specification and 0.732 under the normal. On the other hand, there is little evidence of a systematic relationship between the transition rate and acceptance probabilities estimates. The simple correlations between the τ_c and $\pi_{c(w}c^r)$ are 0.412 under the gamma specification and -0.068 under the normal.[19]

As expected, the estimated level of the acceptance probability is quite sensitive to the family specified; both the ranges and average levels of the acceptance probabilities differ dramatically. Under the gamma, the results have all groups accepting essentially all offers, while half or less appear acceptable under the normal.[20] These differences, in turn, lead to substantial differences in the arrival rate estimates for each group across distributions. Chi-square statistics are reported in the last columns of Tables 4.3 and 4.4 for the test of the overidentifying restriction. Interpreting this as a test of the "goodness-of-fit" for each distribution for each group, the results are slightly more favorable for the normal than the gamma if one does a simple count (three rejections versus four). The distribution parameter estimates are generally precise for both distributions, as well, except for the implied mean offer under the normal distributional assumption--and this appears to be the key to reconciling the difference in the levels of the acceptance probabilities. The normal distributions center roughly at zero for all groups. Since negative offers make no sense, the results can be interpreted as suggesting that the data want to fit themselves to half-normal distributions (i.e., something close to a gamma). Doing a rough normalization for the mass below the mean, the implied acceptance probabilities are about 0.45 to 0.88, numbers that are not quite as far off from the gamma estimates.[21]

To check for sensitivity to measurement error, estimation was also carried out using the first and the second order statistics under each of the offer family specifications. The levels of the acceptance probabilities vary slightly for individual groups (as expected, given slightly different truncation points), but there is little difference in terms of relative fits. As for the key question of interest here, the results for arrivals versus acceptance probabilities are consistent with those reported above.[22]

Section 5 Concluding Remarks

The results presented here suggest that offer arrival rates vary substantially across major demographic groups in the U.S. labor force. They further suggest that variation in the transition rate into employment is directly related to this arrival rate variation. This finding appears robust with respect to specification of a parametric family for the offer distributions. Moreover, little sensitivity to measurement error appears; the results are basically invariant to choice of alternative estimators for the reservation wage.

The estimated level of the acceptance probability for each demographic group does appear quite sensitive to the specification of the offer distribution. Under the gamma specification, all offers appear acceptable for all groups. Under the normal, relatively few offers appear acceptable for any group, although a rough (but more reasonable) interpretation of these results as coming from half-normal distributions also implies acceptance probabilities of about two-thirds to three-quarters. As for the key question of interest here, there is virtually no evidence of a systematic relationship between the transition rate and the acceptance probability under either specification.

Throughout this analysis, the arrival rates and offer distributions are treated as exogenous. Given that the acceptance probability estimates are not far from unity, one might conclude that all variation in transition rates across groups reflects bad luck for some and good luck for others.[23] However, on the basis of my findings alone, we cannot rule out the possibility of variation in search intensity or "systematic search," i.e., that workers apply for jobs that they will almost certainly accept if offered. Under either interpretation, choice on the part of individual workers represents the source of variation in arrival rates and thus transition rates--with or without variation in the acceptance probability. Previous studies that have investigated systematic search hypotheses using data for young workers have generally found that the framework may be relevant empirically (e.g., Jensen and Westergard-Nielsen (1987) and Stern (1989)). The findings reported here suggest that further investigation into systematic search among prime-age and older workers could yield interesting results. More generally, the results suggest the need to address the actual generation of offers, i.e., search technology. At this stage, our understanding in this area is extremely limited.[24]

The remaining assumptions of the model serve primarily to ensure the stationarity of the worker's environment and preferences. It is not clear that structural analysis can be done in a nonstationary framework without extraordinary data or at least some fairly arbitrary assumptions.[25] This remains an area for future research.

Finally, a requirement for using a grouped data approach is the availability of a large representative sample. The data available from the SIPP come closest to satisfying this requirement for the study of dynamics in the U.S. labor market. Further investigation into the potential of both the SIPP data and the usefulness of the grouped data approach using alternative economic models also represents a plan for future research.

<u>References</u>

Baily, Martin N., ed., [1982]: <u>Workers, Jobs, and Inflation</u>, The Brookings Institution, Washington, D.C.

Blau, David M., and Philip K. Robins, [1986]: "Job Search, Wage Offers, and Unemployment Insurance," <u>Journal of Public Economics</u>, vol. 29, 173-197.

Burdett, Kenneth, Nicholas Kiefer, Dale Mortensen, and George Neumann [1984]: "Earnings, Unemployment, and the Allocation of Time Over Time," <u>Review of Economic Studies</u>, vol. 51, 559-578.

Chesher, Andrew and Tony Lancaster, [1983]: "An Econometric Analysis of Reservation Wages," <u>Econometrica</u>, vol. 51, 1661-1176.

Clark, Kim B., and Lawrence H. Summers, [1979]: "Labor Market Dynamics and Unemployment: A Reconsideration," <u>Brooking Papers on Economic Activity</u> 1, 13-72.

Clark, Kim B., and Lawrence H. Summers, [1982]: "Unemployment Insurance and Labor Market Transitions," in M.N. Baily [1982].

Coder, John F., David Burkhead, Angela Feldman-Harkins, and John McNeil, [1987]: "Preliminary Data from the SIPP 1983-84 Longitudinal Research File," Working Paper No. 8702, U.S. Department of Commerce, Bureau of the Census.

Cox, David R. and David Oakes, [1984]: <u>Analysis of Survival Data</u>, Chapman and Hall, London.

David, Martin, ed., [1985]: "The Survey of Income and Program Participation," <u>Journal of Economic and Social Measurement</u>, Special Issue, vol. 13, no. 3-4.

Devine, Theresa J., [1988]: "Interpreting Reemployment Patterns in the Search Framework," Unpublished Ph.D. Dissertation, Cornell University.

Devine, Theresa J. and Nicholas M. Kiefer, [forthcoming]: Empirical Labor Economics in the Search Framework, Oxford University Press.

diFinetti, Bruno, [1975]: Theory of Probability, vol. 2, Wiley, New York.

Ehrenberg, Ronald G., [1981]: "The Demographic Structure of Unemployment Rates and Labor Market Transition Probabilities" in R.G. Ehrenberg, ed., Research in Labor Economics, vol. 3, JAI Press, Greenwich, Connecticut, 241-291.

Fields, Gary S., and George H. Jakubson, [1985]: "Labor Market Analysis Using SIPP," in David, M., ed. [1985].

Flinn, Christopher, and James Heckman, [1982a]: "Models for the Analysis of Labor Force Dynamics," in G. Rhodes and R. Basmenn, eds., Advances in Econometrics, JAI Press, London, CT.

Flinn, Christopher, and James Heckman, [1982b]: "New Methods for Analyzing Structural Models of Labor Force Dynamics," Journal of Econometrics, vol. 18, 115-68.

Galambos, Janos, [1970]: The Asymptotic Theory of Extreme Order Statistics, New York: Wiley.

Hall, Robert, [1972]: "Turnover in the Labor Force," Brooking Papers on Economic Activity, 709-756.

Holzer, Harry, [1988]: "Search Method Use by Unemployed Youth," Journal of Labor Economics, vol. 6, no. 1, 1-20.

Jensen, Peter, and Westergard-Nielsen, Niels, [1987]: "A Search Model Applied to the Transition from Education to Work," Review of Economic Studies, vol. LIV, 461-72.

Kalbfleish, John D. and Robert L. Prentice, [1980]: The Statistical Analysis of Failure Time Data, New York: Wiley.

Lipmann, Steven A., and John J. McCall, [1976]: "The Economics of Job Search, A Survey," Economic Inquiry, vol. 14, 155-189 and 347-368.

Mortensen, Dale T., [1970]: "Job Search, the Duration of Unemployment, and the Phillips Curve," American Economic Review, vol. 60, 505-517.

Mortensen, Dale T., [1986]: "Job Search and Labor Market Analysis," in Ashenfelter and R. Layard, Handbook of Labor Economics, .

Mortensen, Dale T., and George R. Neumann, [1984]: "Choice or Chance? A Structural Interpretation of Individual Labor Market Histories" in G.R. Neumann and N. Westergaard-Nielsen, ed., 98-131.

Narendranathan, Wiji, and Stephen Nickell [1985]: "Modelling the Process of Job Search," Journal of Econometrics, vol. 28, 29-49.

Nelson, Dawn, David B. McMillen, and Daniel Kasprzyk, [1985]: "An Overview of the Survey of Income and Program Participation: Update 1," Working Paper No. 8401, U.S. Department of Commerce, Bureau of the Census.

Neumann, George R., and Neils Westergard-Nielsen, eds., [1984]: Studies in Labor Market Dynamics, Springer-Verlag, Heidelberg.

Perry, George L., [1972]: "Unemployment Flows in the U.S. Labor Market," Brooking Papers on Economic Activity, 245-292.

Phelps, Edmund S., ed., [1970]: Microeconomic Foundations of Employment and Inflation Theory, W.W. Norton and Company, New York.

Ridder, Geert, and Kees Gorter, [1986]: "Unemployment Benefits and Search Behavior: An Empirical Investigation," May, manuscript, Cornell University.

Ryscavage, Paul, and John E. Bregger, [1985]: "New Household Survey and the CPS: A Look at Labor Force Differences," Working Paper No. 8507, U.S. Department of Commerce, Bureau of the Census.

Ryscavage, Paul, and Angela Feldman-Harkins, [1987]: "Work Experience Data from SIPP," Working Paper No. 8703, U.S. Department of Commerce, Bureau of the Census.

Salop, Stephen C., [1973]: "Systematic Job Search and Unemployment," Review of Economic Studies, vol. 41, 191-201.

Sater, Douglas, [1985]: "Enhancing Data from the Survey of Income and Program Participation with Data from Economic Censuses and Surveys," Working Paper No. 8505, U.S. Department of Commerce, Bureau of the Census.

Stern, Steven, [1989]: "Estimating a Simultaneous Search Model," Journal of Labor Economics, vol. 7, no. 3, 348-69.

U.S. Department of Commerce, Bureau of the Census [1984]: Survey of Income and Program Participation User's Guide, 2nd Edition.

U.S. Department of Commerce, Bureau of the Census [1985]: "SIPP Interviewer's Manual."

U.S. Department of Commerce, Bureau of the Census: Survey of Income and Program Participation, 1984 Panel Rectangular Files, Waves 1-4.

U.S. Department of Commerce, Bureau of the Census: Survey of Income and Program Participation, 1984, 1985, and 1986 Panels, Rectangular Core and Topical Modules, Technical Documentation.

Wolpin, Kenneth, [1987]: "Estimating a Structural Search Model: The Transition from School to Work," Econometrica, vol. 55, no. 4, 801-818.

Endnotes

1.I would like to thank Nick Kiefer, George Jakubson, and David Easley for many helpful discussions and comments on this work. Ken Burdett, Alberto Martini, Lars Muus, George Neumann, Geert Ridder, Mark Roberts, and members of the Cornell Labor Workshop are also thanked for comments and suggestions along the way. Remaining errors are mine alone. The data were obtained from ICPSR through CISER at Cornell and most of the computing was done at the Cornell National Supercomputer Facility, funded in part by the National Science Foundation. Support from the Center for Analytic Economics at Cornell is gratefully acknowledged.

An earlier paper, "Interpreting Reemployment Patterns in the 1980s," which used the same estimation approach as that used here was circulated in November 1986 as my job market paper and presented at the December 1986 Winter Meetings of the Econometric Society. Estimation in that paper was carried out using the January 1984 Displaced Worker Survey (DWS), a supplement to the January CPS conducted by the Bureau of the Census in conjunction with the U.S. Department of Labor. The use of the SIPP in the present paper and in Devine (1988) is a consequence of the serious problems discovered with the DPW duration data. (Nonemployment durations are aggregated over an unknown number of spells.) My thanks extend to the BLS staff for bringing these problems with the DPW to my attention and to others who made comments on the approach while I was on the market.

2.Substantial variation in unemployment rates across demographic groups in the U.S., at a given time and across states of the economy, is well documented and has received a fair amount of attention in the economics literature. Perry (1972), Hall (1972), Baily (1982), and Ehrenberg (1981), for example, use CPS gross flow data to calculate the full set of transition rates among unemployment, employment, and nonparticipation that together produce observed variation in unemployment rates across demographic groups. These studies do not investigate the source of variation, however, so the analysis presented here can be viewed as an attempt to go one step farther in studying the demographic composition of the unemployment rate--in that the source of variation in one transition rate is examined. For discussion of reduced form analyses of the roles of labor force participation rates and flows out of employment as sources of variation in the unemployment rates across demographic groups in the U.S., see Devine and Kiefer (forthcoming).

3.This is the same principle underlying the approach taken by Wolpin (1987), Flinn and Heckman (1982), and Stern (1989). However, these studies focus on determining the levels of the arrival rate and the acceptance probability for a single group of workers within the labor force (namely, young white males in the U.S.), as opposed to variation in the roles of these factors across different groups. Their estimation approaches also differ from that employed here at the group level. Lancaster and Chesher (1983) also take a homogeneous population approach to calculating a variety of parameters of interest using some unusual data on reservation wages and other variables for a sample of men in the U.K. All of these papers are reviewed in Devine and Kiefer (forthcoming).

4.The model of individual behavior used here is a simple generalization of the basic models set out in Mortensen (1986) and Lippman and McCall (1976). Note that potential movement here is between the states of unemployment and employment. However, the duration data analyzed below are measurements on spells of being with a job, versus without (defined more carefully below), and the sample is restricted to spells that begin in the sample period. An interpretation of the model as pertaining to movement from nonemployment to employment may be more precise at the empirical stage. The empirical evidence on the importance of distinguishing nonparticipation and unemployment is limited at this time, but it appears that the distinction is perhaps more relevant for adult males than for other groups in the labor force. For discussion of the evidence, see Devine and Kiefer (forthcoming).

5.The probability of receiving at least one offer within a short interval h is thus $\tau_c h + o(h)$, where o(h) is the probability of receiving more than one offer in the interval h and $(o(h)/h) \Rightarrow$ as $h \Rightarrow 0$.

6.Under the Poisson layoff assumption, employment durations for a type c worker also have an exponential distribution (with parameter a_c).

7.The term "structural" here refers to studies that attempt to identify parameters of a tight theoretical structure.

8.Following the (albeit undesirable) convention in the literature, equilibrium considerations are not incorporated into this model.

9.Since the reordering of the individuals in the sample of group c workers does not change the information content of the sample, by assumption (i.e., the subscript i serves only as a label), the individuals are "exchangeable" in the statistical sense. This allows one to treat the sample as if it were an independently and identically distributed sample from some distribution (deFinetti (1975)). This property of each type c sample is exploited at all stages of estimation.

A practical problem with using order statistics for accepted wages is their sensitivity to errors in measurement; all order statistics will be inconsistent in its presence. As a check on my results, I work with the first two order statistics separately, as well as their average. These results are available upon request (Appendix E). I also examine the quality of the reservation wage estimates using a variety of consistency checks (Appendix F).

10.This issue is discussed at length by Ridder and Gorter (1986), Devine (1988), and others.

11.The Pareto family, for example, with density

$$f(w\ ;\ w_{oc},\ \alpha_c) = \alpha w_{oc}^{\alpha} w^{-\alpha-1},\ w \geq w_{oc}\ ,\ \alpha > 0,$$

is excluded on this basis, since the lower bound $w_{oc} \leq w_c^r$ cannot be identified.

12.This estimation procedure is described in greater detail in an appendix that is available upon request (Appendix C).

13.Discussion of the SIPP and the approach taken here to merge the data is provided in an appendix which the author will provide upon request (Appendix D). For additional discussion of SIPP and labor market analysis, see David [1985] and Fields and Jakubson [1985].

14.A restriction to the nonstudent population would have been desirable. Unfortunately, school enrollment information for the 1984 SIPP Panel is quite limited. Only enrollment beyond high school is reported, it is reported for the entire four month reference period preceding a given Wave interview, and there is no indication of the type of education program the person attends. Thus, it is impossible to distinguish high school dropouts from those enrolled in high school or college students from those enrolled in job training programs. No exclusion is used here. Twenty-eight percent of the sample analyzed here reported enrollment at some time in the sixteen month period followed. For both whites and nonwhites, this translates into enrollment of forty to fifty percent of persons less than 25 years of age, about twenty percent for those between 25 to 44, and five to nine percent for older workers. In all cases, these proportions seem too large and suggest ambiguity on the part of respondents. Thses problems with school enrollment data are not present for the 1985 and 1986 Panels of the SIPP.

15.Using both reported hourly wage rates and average hourly earnings, I have observations for 3396 persons. Upon careful inspection of these data, it appeared that there were 34 observations below one dollar simply because of measurement error; I restrict my sample to wage rates greater than or equal to one dollar.

16.This is distinct from the standard BLS definition of unemployment (i.e., no employment arrangement with an employer and actively seeking such, or being on definite or temporary layoff). It is also distinct from being out of work (which may include absence due to illness, a labor dispute, or vacation). The SIPP data for labor market status do allow one to use standard definitions, but not without a substantial amount of human effort and computer time if one wishes to link accepted wages with the weekly labor market activity data. The problems with linking arise because labor market income data are reported for up to two employers in the SIPP in a separate section of the survey from the weekly labor market activity data, where the employers discussed are the two most recent <u>or</u> those for whom the most hours were worked. Dates of employment within the four month Wave sample period are provided in the income section for each of the two employers, but the criteria used for employer selection make direct use of the employment date data problematic because of overlapping job spells. Meanwhile, the order must be sorted out to make the link with the weekly activity data. All of these considerations lead to use of the somewhat aggregated job versus no job durations in the present first round analysis of the SIPP. I am currently engaged in constructing more precise work week-by-week histories and will exploit these data in future work.

17.The employment duration data described in Table 3.1 are for those who experienced initialized job spells and initialized joblessness spells. When the second criterion is not imposed, the sample mean for the uncensored job durations is 19.17 weeks.

18.Corresponding estimates of the distribution parameters θ_c and their standard errors are presented in Appendix 1. Note that the sample sizes for all nonwhite groups are significantly smaller than sample sizes for the white groups. This may explain the higher estimates for the reservation wage; the positive bias of the order statistics can be shown to be decreasing with sample size. Note that this will decrease the acceptance probability estimate and thus push up the estimates for the arrival rates. Consequently, the white-nonwhite differential in the arrival rates is probably understated by the numbers presented here.

19. Regressions of τ_c on δ_c yield coefficients of 0.995 (s.e. 0.015) under the gamma with all group observations included and 0.993 (s.e. 0.012) when only the twelve observations with χ^2 less than 5.02 are included. For the normal, regressions of τ_c on δ_c yield coefficients of 0.183 (s.e. 0.046) with all group observations included and 0.205 (s.e. 0.038) when only the thirteen observations with χ^2 less than 5.02 are included. On the other hand, regressions of τ_c on $\pi_{c(w}c^r)$ yield coefficients of 0.309 (s.e. 0.183) under the gamma with all group observations included and 0.645 (s.e. 0.338) when only the twelve observations with χ^2 less than 5.02 are included. For the normal, regressions of τ_c on $2(\pi_{c(w}c^r))$ yield coefficients of -0.007 (s.e. 0.027) with all group observations included and -0.051 (s.e. 0.036) when only the thirteen observations with χ^2 less than 5.02 are included; the use of 2π is based on the half-normal interpretation of the results discussed in the text.

20.The limited variation in the acceptance probability across groups should not be confused with a lack of variation across groups in the offer distributions faced. Quite the contrary, as indicated by the sample mean accepted wages $E_c[w|w\geq w_c^r]$, the offer distributions vary substantially.

21.Wolpin [1987] reports that he obtains negative estimates for mean offers when he attempts to fit a normal distribution to wage data for young male workers. His results might be interpreted as suggesting half-normal distributions as well.

22.These results are available upon request (Appendix E). Also, experimentation with lognormal distributions yielded estimates of unity for the acceptance probability.

23.In looking at the roles of arrivals versus acceptance probabilities in producing variation across individuals, Mortensen and Neumann (1984) maintain this assumption and accordingly describe the two factors as "chance" and "choice," respectively.

24.Holzer (1988), for example, provides some descriptive evidence on the search process of low-income youth in the U.S. For a discussion of work in this area to date, see Devine and Kiefer (forthcoming), Chapter 7.

25.Wolpin (1987), for example, works with data for young U.S. males and relaxes stationarity by imposing a finite search horizon, defined as a date after which all offers are accepted.

Table 3.1 DATA SUMMARY

Variable	N	Mean	Standard Deviation
Race (White = 1)	5214	0.84	
Sex (Male = 1)	5214	0.47	
Marital Status (Married = 1)	5214	0.45	
Age	5214	32.46	13.64
Education (Years)	5214	12.72	3.08
Durations: (Weeks) Without Job	5214	14.96	13.74
Uncensored	4112	9.83	8.14
With Job	3907	13.16	11.30
Uncensored	3139	10.82	9.65
Accepted Wage (Hourly Wage or Average Hourly Earnings)	3396	5.88	4.37
Accepted Job Weekly Hours	3521	35.69	12.39

Table 4.1 PARTITION BY DEMOGRAPHIC CHARACTERISTICS

Characteristic	Group Definition			
Age	16-19	20-24	25-44	45-64
Race	White		Nonwhite	
Sex	Male		Female	

Table 4.2 GROUP TRANSITION RATES

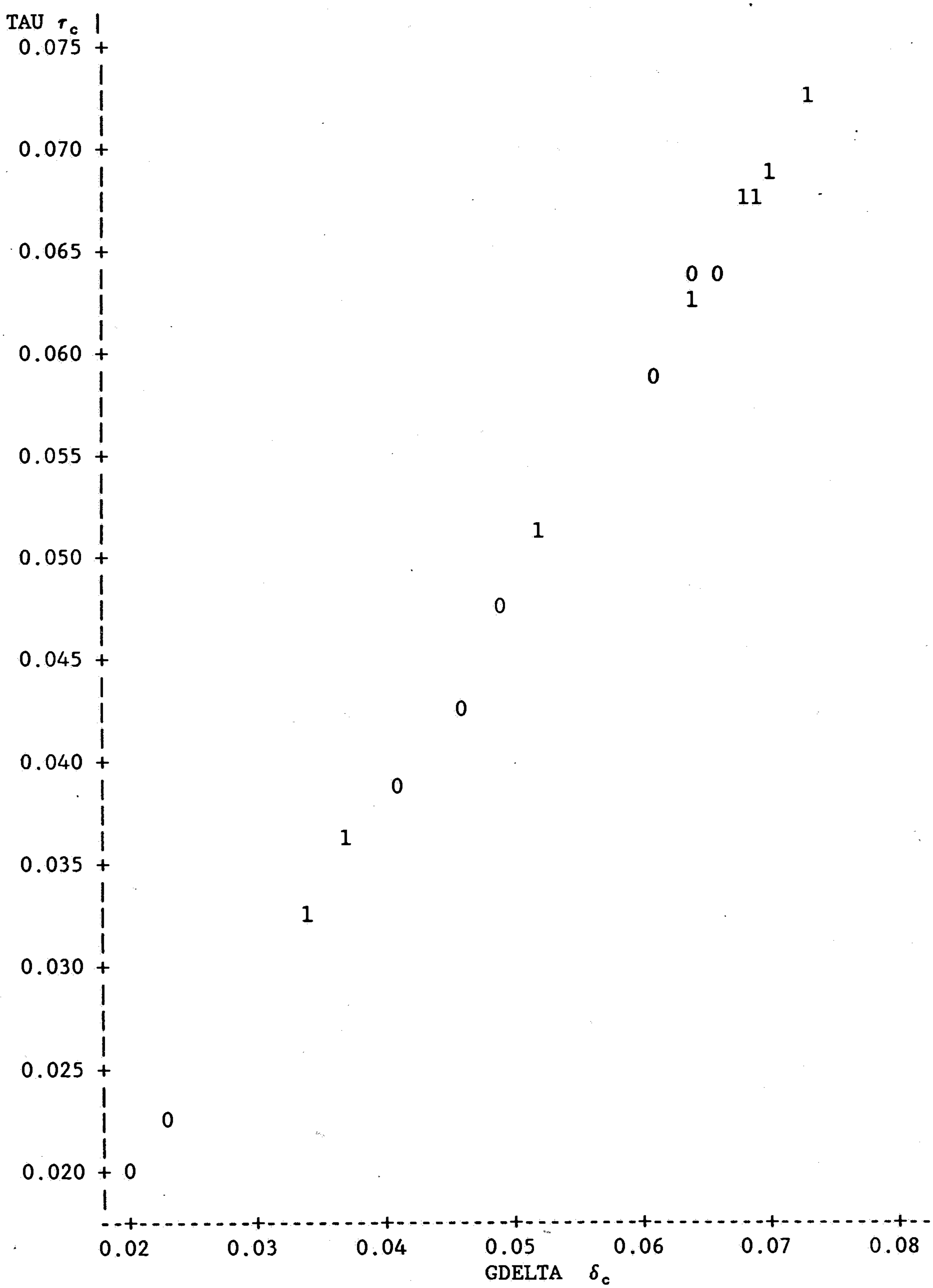
GROUP:	N WAGE	N DUR	TAU τ_c	E(DUR)
White Males:				
Ages 16-19	296	360	0.073	13.69
Ages 20-24	404	516	0.069	14.49
Ages 25-44	556	801	0.068	14.70
Ages 45-64	187	392	0.036	27.77
Nonwhite Males:				
Ages 16-19	44	57	0.059	16.94
Ages 20-24	66	88	0.064	15.62
Ages 25-44	89	150	0.048	20.83
Ages 45-64	25	73	0.020	50.00
White Females:				
Ages 16-19	240	293	0.067	14.92
Ages 20-24	363	527	0.063	15.87
Ages 25-44	598	1005	0.051	19.60
Ages 45-64	226	484	0.033	30.30
Nonwhite Females:				
Ages 16-19	40	47	0.064	15.62
Ages 20-24	63	103	0.042	23.80
Ages 25-44	129	233	0.039	25.64
Ages 45-64	36	85	0.023	43.47

Table 4.3 GAMMA DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS

GROUP:	N WAGE	WR2A w_c^r	WAGE $E(w \mid w \geq w_c^r)$	TAU τ_c	GPI $\pi_c(w_c^r)$	GDELTA δ_c	GCHI[a] χ^2
White Males:							
Ages 16-19	296	1.054	4.067	0.073	0.999	0.073	4.152
Ages 20-24	404	1.054	5.640	0.069	0.995	0.070	1.763
Ages 25-44	556	1.684	8.261	0.068	0.980	0.069	1.097
Ages 45-64	187	1.364	9.857	0.036	0.984	0.037	1.081
Nonwhite Males:							
Ages 16-19	44	2.249	3.667	0.059	0.956	0.061	2.132
Ages 20-24	66	1.169	4.445	0.064	0.997	0.064	3.646
Ages 25-44	89	1.634	6.849	0.048	0.966	0.049	2.518
Ages 45-64	25	2.464	8.739	0.020	0.968	0.020	1.974
White Females:							
Ages 16-19	240	1.054	3.792	0.067	0.998	0.068	1.038
Ages 20-24	363	1.064	4.493	0.063	0.990	0.064	1.480
Ages 25-44	598	1.029	5.842	0.051	0.984	0.052	5.518
Ages 45-64	226	1.539	6.043	0.033	0.968	0.034	7.269
Nonwhite Females:							
Ages 16-19	40	2.125	3.602	0.064	0.974	0.066	9.295
Ages 20-24	63	3.059	4.355	0.042	0.906	0.046	15.588
Ages 25-44	129	1.484	5.911	0.039	0.964	0.041	1.166
Ages 45-64	36	1.579	4.655	0.023	0.982	0.023	1.102

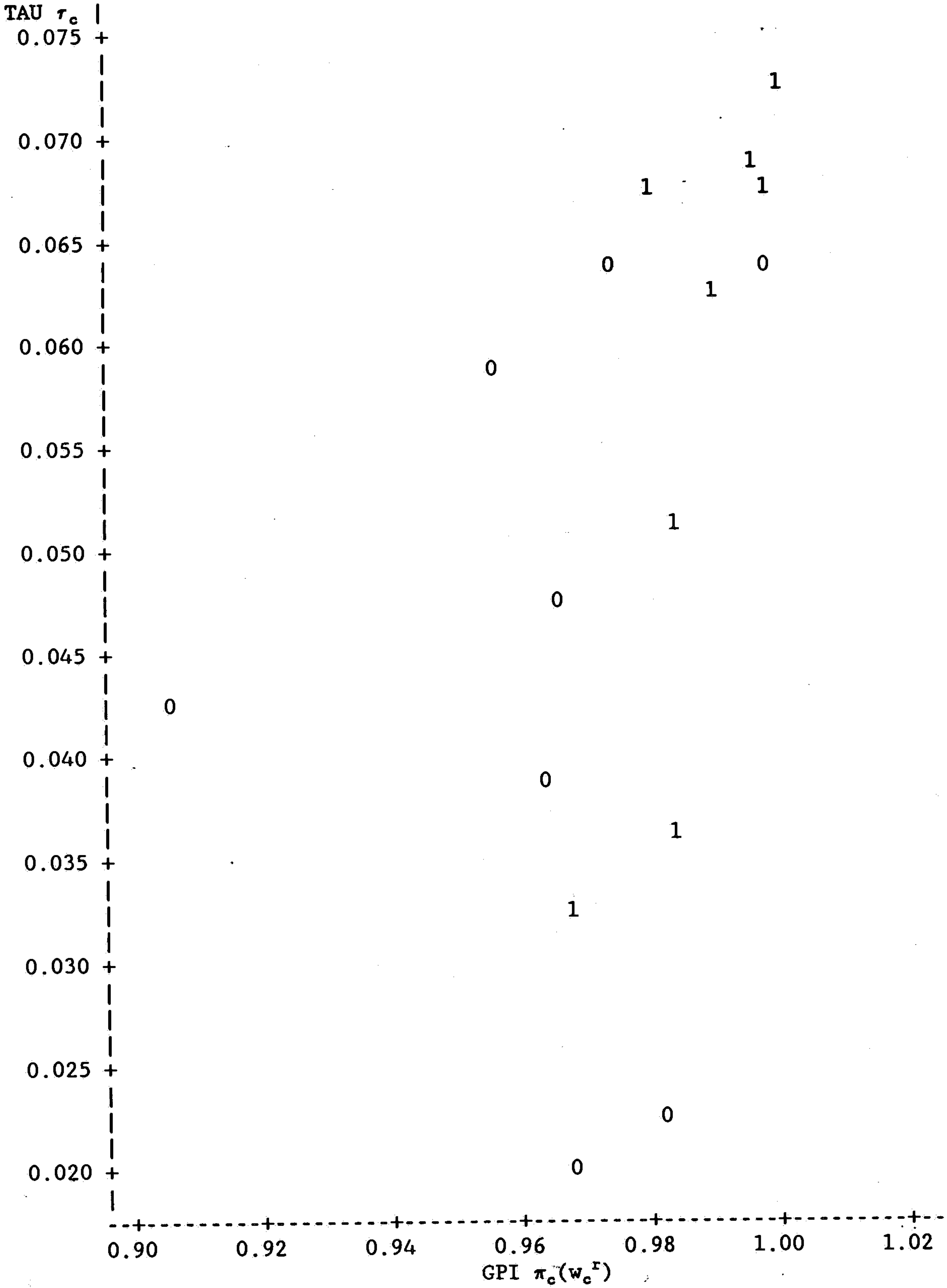
[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

Figure 4.1 GAMMA DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS



PLOT OF TAU*GDELTA SYMBOL IS RACE (WHITE=1)

Figure 4.2 GAMMA DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS



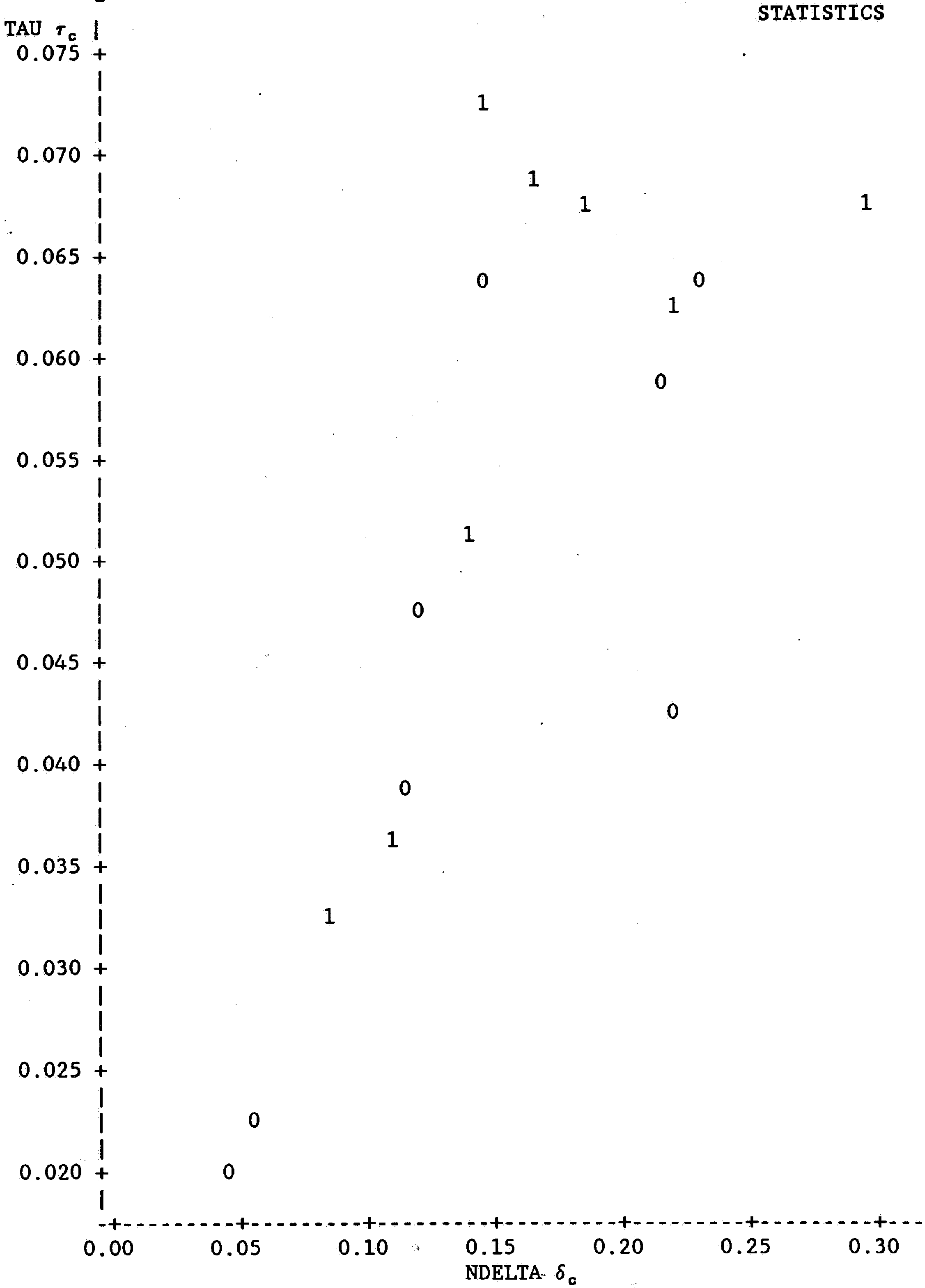
PLOT OF TAU*GPI SYMBOL IS RACE (WHITE=1)

Table 4.4 NORMAL DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS

GROUP:	N WAGE	WR2A w_c^r	WAGE $E(w \mid w \geq w_c^r)$	TAU τ_c	NPI $\pi_c(w_c^r)$	NDELTA δ_c	NCHI[a] χ^2
White Males:							
Ages 16-19	296	1.054	4.067	0.073	0.504	0.145	30.927
Ages 20-24	404	1.054	5.640	0.069	0.414	0.167	1.750
Ages 25-44	556	1.684	8.261	0.068	0.367	0.185	0.078
Ages 45-64	187	1.364	9.857	0.036	0.325	0.112	0.414
Nonwhite Males:							
Ages 16-19	44	2.249	3.667	0.059	0.272	0.217	0.717
Ages 20-24	66	1.169	4.445	0.064	0.437	0.146	0.577
Ages 25-44	89	1.634	6.849	0.048	0.395	0.121	0.293
Ages 45-64	25	2.464	8.739	0.020	0.430	0.046	0.100
White Females:							
Ages 16-19	240	1.054	3.792	0.067	0.229	0.296	0.369
Ages 20-24	363	1.064	4.493	0.063	0.290	0.219	0.031
Ages 25-44	598	1.029	5.842	0.051	0.367	0.139	0.179
Ages 45-64	226	1.539	6.043	0.033	0.398	0.083	34.011
Nonwhite Females:							
Ages 16-19	40	2.125	3.602	0.064	0.283	0.228	0.809
Ages 20-24	63	3.059	4.355	0.042	0.189	0.222	11.677
Ages 25-44	129	1.484	5.911	0.039	0.348	0.114	0.048
Ages 45-64	36	1.579	4.655	0.023	0.408	0.056	0.311

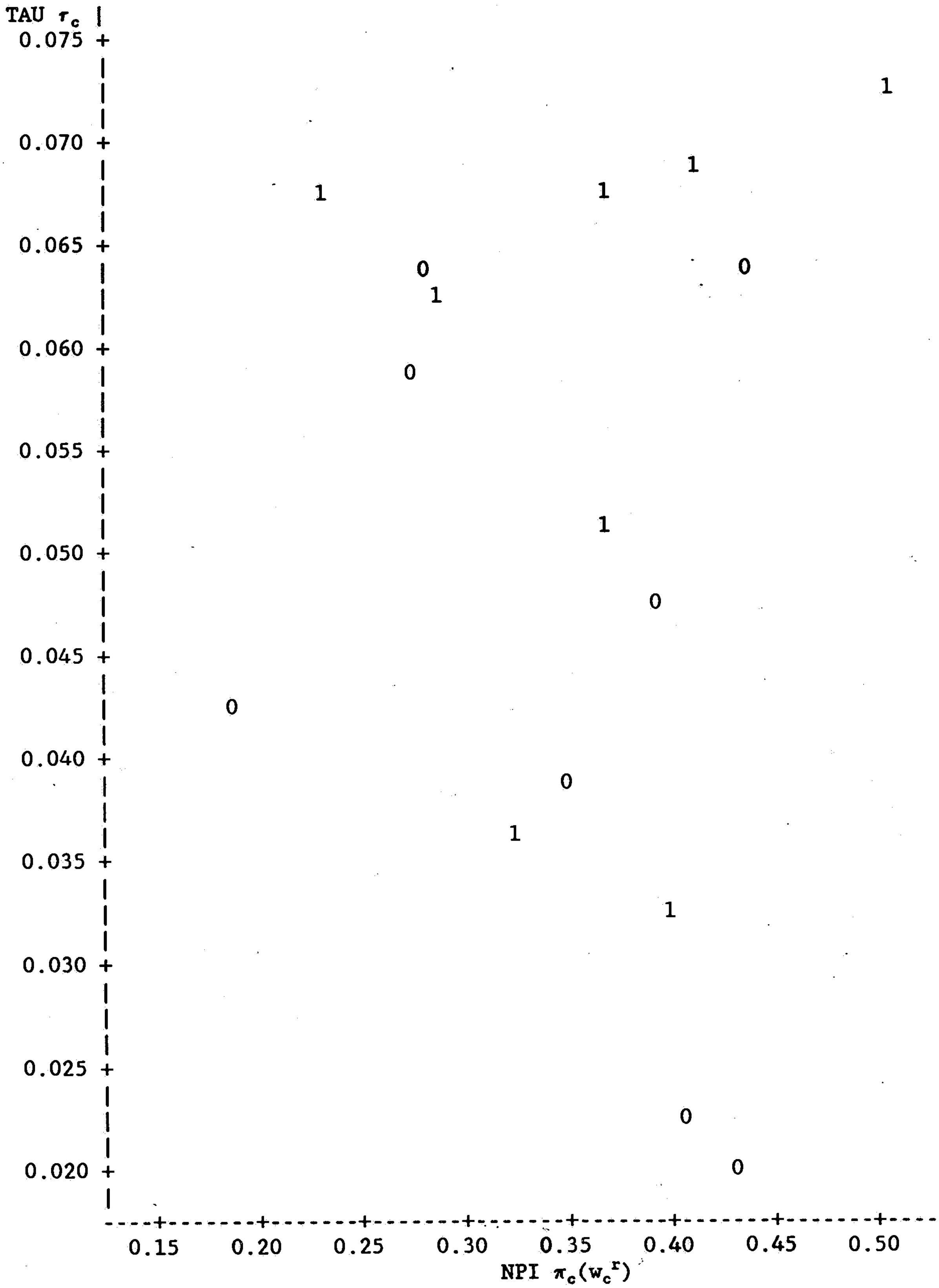
[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

Figure 4.3 NORMAL DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS



PLOT OF TAU*NDELTA SYMBOL IS RACE (WHITE=1)

Figure 4.4 NORMAL DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS



PLOT OF TAU*NPI SYMBOL IS RACE (WHITE=1)

Appendix 1

Parameter Estimates for the Offer Distribution

Table A1 presents parameter estimates for the gamma and normal distributions using the average of the first order two statistics as the estimator for the reservation wage, corresponding to the search parameter estimates reported in Section 4. The implied mean and variance for the gamma are also reported (GMU and GVAR, respectively).

The size of the maximum of the absolute value of the gradient at the reported estimates is reported along with the parameter estimates and their asymptotic standard errors. For some groups, the parameters did not converge under the gradient convergence criterion of 10^{-4}. However, experimentation with alternative starting values and alternative step sizes failed to produce any change in the parameter values from those reported for these cases and the distance appeared to be at a minimum.

The standard errors for the parameter estimates are quite small in statistical terms for both distributions, with the exception of the mean of the offer distribution under the normal specification. As for the relative fits of the two distributions, the values of chi-square statistic are only slightly more likely to lead to a rejection of the overidentifying restriction under the gamma than under the normal.

Obviously, variation in sample size across groups should be considered in interpreting the results. While the overall sample size is quite large, the sample sizes for some of the nonwhite groups are quite small. By the same token, it appears that no alternative data source provides larger samples for these groups, while providing observations on the variables required.

Table A.1 GAMMA DISTRIBUTION: WR2A – AVERAGE OF FIRST TWO ORDER STATISTICS

PARAMETER ESTIMATES (Standard Errors in Parentheses)

GROUP:	N WAGE	ALPHA α_c	BETA β_c	GMU $E_c(w)$	GVAR Var(w)	GCHI[a] χ^2	MAX F[b]
White Males:							
Ages 16-19	296	14.253	0.279	3.978	1.110	4.152	0.000
			(1.713)	(0.035)			
Ages 20-24	404	4.471	1.234	5.520	6.817	1.763	0.000
			(0.389)	(0.121)			
Ages 25-44	556	3.338	2.400	8.014	19.236	1.097	0.000
			(0.230)	(0.175)			
Ages 46-64	187	2.664	3.457	9.211	31.846	1.081	0.000
			(0.407)	(0.564)			
Nonwhite Males:							
Ages 16-19	44	17.835	0.199	3.553	0.707	2.132	0.647
			(1.171)	(0.013)			
Ages 20-24	66	7.997	0.508	4.064	2.066	3.646	0.003
			(2.287)	(0.151)			
Ages 25-44	89	3.331	1.908	6.357	12.130	2.518	0.001
			(0.693)	(0.424)			
Ages 45-64	25	3.992	2.100	8.387	17.621	1.974	0.002
			(1.554)	(0.737)			
White Females:							
Ages 16-19	240	9.101	0.402	3.662	1.473	1.038	0.000
			(1.616)	(0.073)			
Ages 20-24	363	4.882	0.886	4.329	3.839	1.480	0.000
			(0.873)	(0.165)			
Ages 25-44	598	3.221	1.716	5.527	9.485	5.518	0.000
			(0.273)	(0.160)			
Ages 45-64	226	3.756	1.472	5.532	8.147	7.269	0.000
			(0.502)	(0.210)			
Nonwhite Females:							
Ages 16-19	40	18.999	0.185	3.519	0.652	9.295	6.184
			(2.886)	(0.026)			
Ages 20-24	63	41.500	0.091	3.817	0.351	15.588	21.447
			(30.930)	(0.121)			
Ages 25-44	129	3.427	1.615	5.535	8.940	1.166	0.000
			(0.525)	(0.263)			
Ages 45-64	36	6.578	0.671	4.416	2.965	1.102	0.000
			(3.273)	(0.328)			

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

[b]MAX F is the maximum absolute value of the gradient at the reported estimates. In some cases, this value exceeds the desired value for convergence. In such cases, the sum of squares and reported parameter values were unchanging.

Table A.2 NORMAL DISTRIBUTION: WR2A - AVERAGE OF FIRST TWO ORDER STATISTICS

PARAMETER ESTIMATES (Standard Errors in Parentheses)

GROUP:	N WAGE	NMU μ_c	NSIG σ_c	NCHI[a] χ^2	MAX F[b]
White Males:					
Ages 16-19	296	1.098	4.252	30.927	0.022
		(0.054)	(0.010)		
Ages 20-24	404	-0.361	6.591	1.750	0.003
		(4.379)	(1.951)		
Ages 25-44	556	-1.576	9.642	0.078	0.000
		(0.201)	(0.103)		
Ages 45-64	187	-4.301	12.546	0.414	0.015
		(0.061)	(0.413)		
Nonwhite Males:					
Ages 16-19	44	0.751	2.475	0.717	0.000
		(0.149)	(0.016)		
Ages 20-24	66	0.430	4.669	0.577	0.009
		(1.053)	(0.382)		
Ages 25-44	89	-0.318	7.353	0.293	0.142
		(8.623)	(3.867)		
Ages 45-64	25	0.910	8.865	0.100	0.000
		(0.841)	(0.074)		
White Females:					
Ages 16-19	240	-2.432	4.710	0.369	0.185
		(0.034)	(0.077)		
Ages 20-24	363	-1.950	5.456	0.031	0.008
		(0.083)	(0.078)		
Ages 25-44	598	-1.316	6.923	0.179	0.000
		(0.152)	(0.053)		
Ages 45-64	226	-0.073	6.290	34.011	1.116
		(1.697)	(0.840)		
Nonwhite Females:					
Ages 16-19	40	0.665	2.549	0.809	0.010
		(0.262)	(0.051)		
Ages 20-24	63	0.925	2.425	11.677	0.755
		(0.311)	(0.023)		
Ages 25-44	129	-1.041	6.496	0.048	0.000
		(0.535)	(0.050)		
Ages 45-64	36	0.533	4.526	0.311	0.005
		(1.023)	(0.302)		

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

[b]MAX F is the maximum absolute value of the gradient at the reported estimates. In some cases, this value exceeds the desired value for convergence. In such cases, the sum of squares and reported parameter values were unchanging.

Appendix A

The Reservation Wage

The reservation wage is the wage that equates the expected present value of employment at that wage $V_c^e(w)$ and the expected present value of continued optimal search V_c^u for a worker of type c, i.e.,

(A.1) $V_c^e(w_c^r) = V_c^u$.[1]

The time-invariance of the policy follows from the stationarity of the worker's environment defined in the text.

More formally, since the net income flow while unemployed b_c is a constant, since offers are independent and identically distributed, and since the distribution $F_c(w)$ and arrival rate ${}_c$ are known and time-invariant, the value of optimal search V_c^u is a constant defined implicitly by the equation

$$\text{(A.2)} \quad V_c^u = \frac{1}{1 + r_c h} b_c h + \frac{\delta_c h}{1 + r_c h} E_w[\max\{V_c^e(w), V_c^u\}]$$

$$+ (1-\delta_c h) \frac{1}{1 + r_c h} V_c^u + o(h),$$

where h denotes a short period of time. The first term on the right side of this equation is the discounted present value of net unemployment income over the interval h. The second is the probability of receiving an offer in the interval h times the discounted expected value of following the optimal policy once an

---

[1]Bellman's optimality principle asserts that a worker's current choice maximizes the sum of the flow of utility in the current period and the mathematical expectation of the worker's discounted expected flow of utility over the future, given that all future decisions will be made optimally. Application of the principle requires that a worker's preferences over the future can be taken to be the discounted sum of returns accruing over the future. The maximization of discounted expected lifetime income satisfies this requirement.

offer w is received. The third term is the probability of no offer in the interval h times the discounted value of optimal search thereafter.

The expected present value of accepting a given wage offer w in this model, $V_c^e(w)$, does not depend on when the offer is received. It is defined by

$$(A.3)\quad V_c^e(w) = \frac{1}{1 + r_c h}\, wh + \frac{1}{1 + r_c h}\,[(1-a_c h)V_c^e(w) + a_c h V_c^u] + o(h).$$

The first term on the right is the present value of income which will be received over a short interval of length h. The second is the present value of expected lifetime income at the end of this period. This is simply a weighted average of the worker's expected income if he or she remains employed at w and expected income if he or she is laid off and searches thereafter. The weights are simply the probabilities of each of these events occurring over the interval h.

Solving (A.3) for $V_c^e(w)$ and letting h --> 0 yields

$$(A.4)\qquad V_c^e(w) = \frac{w}{r_c + a_c} + \frac{a_c V_c^u}{r_c + a_c},$$

which is continuous and strictly increasing in w. It follows that the wealth maximizing and therefore optimal

policy is a reservation wage policy: accept any wage w such that $w \geq w_c^r$, where the reservation wage w_c^r solves

$$(A.5)\qquad V_c^e(w_c^r) = \frac{w_c^r}{r_c + a_c} + \frac{a_c V_c^u}{r_c + a_c} = V_c^u$$

or, equivalently,

$$(A.6)\qquad \frac{w_c^r}{r_c} = V_c^u .$$

Substitution of (A.5) for $V_c^e(w)$ in (A.2) and using (A.6) to eliminate V_c^u yields the fundamental reservation wage equation for this model

$$\text{(A.7)} \quad \frac{w_c^r}{r_c} = \frac{1}{1 + r_c h} b_c h + \frac{\delta_c h}{1 + r_c h}$$

$$\cdot E_w\left[\max\left\{\left[\frac{w}{r_c + a_c} + \frac{a_c w_c^r/r_c}{r_c + a_c}, \frac{w_c^r}{r_c}\right]\right]\right.$$

$$+ \frac{(1-\delta_c h)}{1 + r_c h}\frac{w_c^r}{r_c} + o(h).$$

Passing to the limit, this becomes

$$\text{(A.8)} \quad w_c^r = bi + \frac{\delta_c}{r_c + a_c}\int_{w_c^r}^{\infty} (w_c - w_c^r)\, dF_c(w_c).$$

By evaluating the integral in (A.8), rearranging terms, and letting $r_c^* = r_c + a_c$ denote a discount rate which accounts for the probability of layoff as well as the individual's rate of time preference, this condition may be rewritten in a form which more readily affords an intuitive interpretation of w_c^r:

$$\text{(A.9)} \quad (w_c^r - b_c)r_c^* = \{\ (E_w[w|w \geq w_c^r] - w_c^r)$$

$$\cdot ([1 - F_c(w_c^r)]\delta_c)\ \}.$$

The lefthand side of (A.9) gives the marginal cost of rejecting an offer equal to w_c^r and continuing to search. This is simply the imputed interest income flow on the difference between incomes in the two alternatives. The righthand side gives the marginal expected gain in future earnings from continued search, given that an offer will be accepted only if it exceeds the reservation wage, times the instantaneous probability that an acceptable offer will be received. That is, the righthand side gives the expected marginal return to continued optimal search. The reservation wage which represents the optimal policy for the worker is thus simply the wage rate which equates the marginal cost and marginal benefit of search activity.

Appendix B

Consistency of the Order Statistics

Proving the consistency of the k^{th} order statistic for w_c^r is a straightforward exercise. I sketch a proof here, omitting all subscripts c for convenience.[1]

Let $W_1^n \le W_2^n \le W_3^n \ldots \le W_n^n$ denote the order statistics from the distribution function for accepted wages $F(w|w > w^r)$ for a random sample of size n. The marginal distribution function for W_k^n, $k = 1, 2, \ldots n$, is given by

$$(B.1) \quad G_k^n(w) = \sum_{j=k}^{n} \frac{n!}{(n-j)!j!} [F(w|w \ge w^r)]^j [1-F(w|w \ge w^r)]^{n-j}, \quad w \ge w^r$$

$$= 0 \text{ , otherwise.}$$

Now define the sequence of nonnegative random variables

$$(B.2) \quad Y_k^n = W_k^n - w^r \text{ .}$$

Consistency of W_k^n for w^r requires that $\lim Pr(Y_k^n \le y) = 1$ for all $y > 0$. From B.1 and B.2, we have that

$$(B.3) \quad H_k^n(y) = P(Y_k^n \le y) = Pr(W_k^n \le y + w^r) = G_k^n(y + w^r)$$

$$= 1 - \sum_{j=0}^{k-1} \frac{n!}{(n-j)!j!} [F(y + w^r|w > w^r)]^j$$

$$\cdot [1-F(y + w^r|w) > w^r)]^{n-j}, \quad y > 0.$$

Since $F(\cdot)$ is a distribution function, it lies in the unit interval so that the sum in this equation tends to zero as n grows large and

$$(B.4) \quad \lim_{n \to \infty} H_k^n(y) = 1 \text{ ,}$$

as required for consistency of W_k^n.

---

[1]See Galambos (1978) for a complete treatment of the properties of order statistics, including the asymptotic distributions under some alternative distributional assumptions on w.

Appendix C

Estimation of the Offer Distribution Parameters

Consider a homogeneous population. Let $F(w|\theta)$ and $f(w|\theta)$ denote the distribution function (df) and density, respectively, for the wage distribution from which individuals in this population draw their offers. I specify this distribution up to the p x 1 vector of parameters θ. Let w^r denote the reservation wage for all individuals in this population. The density and df for accepted wages corresponding to this distribution are then defined as

$$\text{(C.1)}\quad f^T(w|w \geq w^r, \theta) = \frac{f(w|\theta)}{1 - F(w^r|\theta)} \quad , \ w \geq w^r$$

$$= 0 \qquad , \text{ otherwise}$$

and

$$\text{(C.2)}\quad F^T(w|w \geq w^r, \Theta) = \int_{w^r}^{w} \frac{f(u|\theta)}{1 - F(w^r|\theta)} du \quad , \ u \geq w^r$$

$$= 0 \qquad , \text{ otherwise.}$$

My data consist of observations on the accepted wages W_i for a size n sample of workers. My objective is to estimate the vector θ using this accepted wage data.[1] I do this using the following three stage procedure.

Assume that the first 2k moments $E(W^j|W \geq w^r)$ of the truncated distribution $F^T(w|w \geq w^r, \theta)$ exist. The first stage of my estimation procedure is estimation of these moments. These moments are functions of the

[1] Throughout this section, I treat w_c^r as if it were a known parameter. In practice, I use an estimator for w_c^r which converges faster than my estimator for θ_c (rate $N_c^{1/2}$, as compared to N_c, so that the asymptotic properties of my estimates for θ_c are not affected.

vector θ, so I denote them by

$m_j(\theta)$, j = 1,2,...2k.

Let $S_{jn} = (1/n)\Sigma W^j{}_i$ denote the j^{th} sample moment, j = 1,2,...2k, for a sample of size n. I use S_{jn} as the first stage estimator, i.e., $m_j^*(\theta) = S_{jn}$. Assuming that the sample is independently and identically distributed (iid), I have that

(i) (Kolmogorov's Strong Law of Large Numbers) $S_{jn} \Rightarrow m_j(\theta)$, almost surely, as $n \Rightarrow \infty$.

S_{jn} thus provides a strongly consistent estimate for the j^{th} moment $m_j(\theta)$), j = 1,2,...2k. Also, letting $S_n = (S_{1n}, S_{2n}, \ldots S_{kn})'$ denote the vector of the first k sample moments and $m(\theta) = (m_1(\theta), \ldots m_j(\theta), \ldots m_k(\theta))'$ denote the vector of the first k moments of the truncated distribution $F^T(w|w \geq w^r, \theta)$, under the usual regularity conditions and the iid assumption for my sample, I have

(ii) (Central Limit Theorem) $\sqrt{n}(S_n - m(\theta)) \xrightarrow{d} N(\underline{0}, \Omega)$, where $\underline{0}$ denotes a kx1 vector of zeros and $\Omega = [\sigma_{mj}]$, a kxk covariance matrix with $\sigma_{mj} = Cov(W^m, W^j)$, j,m = 1,2,...k.

The second stage of the estimation procedure is calculation of a consistent estimate Ω^* of the matrix Ω. Under the iid assumption, the elements of this matrix are defined as

(C.3) $$\sigma_{mj} = E(W^m W^j) - E(W^m)E(W^j) = q(E(W^j)),$$

q a continuous function. Thus, given that the first 2k moments of the truncated offer distribution $F^T(w|w \geq w^r, \theta)$ exist, I have by way of appeal to the iid assumption, the strong law of large numbers (i) above, and Slutsky's Theorem,

(iii) If $q(\cdot)$ is continuous and $X_n \xrightarrow{P} X$, then

$$q(X_n) \xrightarrow{P} q(X),$$

that consistent estimates for the elements of the matrix Ω are provided by

(C.4) $\sigma_{mj}^* = S_{m+j} - S_m S_j$

since convergence in probability is implied by almost sure convergence. [2]

The model at the third stage of estimation may be written as

(C.5) $S_n = m(\theta)$,

a system of k ($\geq$ p, the dimension of θ) nonlinear simultaneous equations in θ. To estimate the vector θ, I solve the minimum distance problem

$$\text{(C.6)} \quad \min_{\substack{\theta \\ \theta \in \Theta}} \; D(\theta) = [S_n - m(\theta)]' \, A \, [S_n - m(\theta)] \; ,$$

where the matrix A is the metric with which we measure distance and the space over which $D(\cdot)$ is minimized depends on the specification of $F(w|\theta)$. The efficient metric is the inverse of the asymptotic covariance matrix for S_n, Ω^{-1}. Using the second stage consistent estimate Ω^* for Ω does not affect the asymptotics.

The first order conditions for this problem may be written as

$$\text{(C.7)} \quad d(\theta) = \frac{\partial\, D(\theta)}{\partial\, \theta} = \frac{\partial\, m(\theta)'}{\partial\, \theta} \, \Omega^{*-1} \, [S_n - m(\theta)] = 0 \; .$$

Assuming that the moments $m(\theta)$ are continuously differentiable, these equations are well defined. The minimum distance estimator θ^* is obtained by

[2]See Rao (p.124) for a proof of the Lemma.

solving equations (C.7) for θ.[3] Letting θ^* denote the convergent estimate for θ, it can be shown that θ^* is consistent and, also, that $\sqrt{n}(\theta^* - \theta)$ converges in distribution to $N(\underline{0}, V(\theta))$ where the matrix $V(\theta)$ is defined as

$$\text{(C.8)} \quad V(\theta) = \left[\frac{\partial\ m(\theta)}{\partial\ \theta}\ \Omega^{-1}\ \frac{\partial\ m(\theta)}{\partial\ \theta'}\right]^{-1}.$$

$[(1/n)V(\theta^*)]$ provides a consistent estimate of this matrix.[4]

If the specification of the offer distribution is correct, $nD(\theta^*)$ is asymptotically distributed chi-square with k-p degrees of freedom, the number of overidentifying restrictions in the model (C.6). This provides me with a "goodness-of-fit" test.

For each specification of the offer distribution $F(w|\theta_c)$, the above procedure is carried out for each of the groups c, c = 1,2,...C in a given partition of the total population, since each of these groups is assumed to represent a distinct homogeneous population characterized by a set of parameters $\{\tau_c, \delta_c, \pi_c(w_c^r)\}$. The two distributions $F(w|\theta_c)$ with which I

---

[3] Minimization is carried out using the Gauss-Newton method (programmed in Proc Matrix (IML) in SAS). At the j+1st iteration, my estimates are thus

$$\theta^{*j+1} = \theta^{*j} + kH^{-1}(\theta^{*j})d(\theta^{*j}),$$

where k is a scalar and the matrix $H(\theta^{*j})$ is defined as

$$H(\theta^{*j}) = \frac{\partial\ m(\theta^{*j})}{\partial\ \theta}\ \Omega^{*-1}\ \frac{\partial\ m(\theta^{*j})}{\partial y\ \theta'}.$$

Analytical derivatives are calculated at each iteration, with a numerical derivative for the digamma where required. The criterion for convergence is $\max(d(\theta^{*j})) < 10^{-4}$.

[4] These proofs may be found in Chow (1983, Chapter 7) which, except for the proof of consistency, is based on Malinvaud (1970).

primarily work are the normal and gamma. Each is a two parameter family. I work with the first three moments of the truncated distributions so that I have a "goodness of fit" test in the form of a test of one overidentifying restriction. The densities $f(w|\theta_c)$ and $f^T(w|w \geq w_c^r)$, vectors θ_c, and 3x1 vector of theoretical moments $m(\theta_c)$ for each of these distributions are as follows (The subscript c is omitted here for convenience.):

Normal

Let offers W be distributed $N(\mu,\sigma)$. The vector θ is defined as $(\mu, \sigma)'$, $-\infty < \mu < \infty$ and $\sigma > 0$. The density for W may be written as

$$\text{(C.9)} \quad f(w|\mu, \sigma) = (2\ \sigma)^{-\sigma} \exp\{-(w-\mu)/2\sigma\}, \qquad -\infty < w < \infty,$$

and the density for accepted wages, given the reservation wage w^r, is

$$\text{(C.10)} \quad f^T(w|w \geq w^r, \mu, \sigma) = \frac{\exp\{-(w-\mu)/2\sigma\}}{\int_{w^r}^{\infty} \exp\{-(w-\mu)/2\sigma\}dw},$$

for $w \geq w^r$. Letting $\Phi(\cdot)$ denote the standard normal df, $\phi(\cdot)$ denote the standard normal density, and defining the function $Q(\mu,\sigma)$

$$\text{(C.11)} \quad Q(\mu,\sigma) = \frac{\phi((w^r - \mu)/\sigma)}{1 - \Phi((w^r - \mu)/\sigma)},$$

the moments for the truncated distribution may be written as

$$\text{(C.12)} \quad \text{(a)} \quad m_1(\mu,\sigma^2) = \mu + Q(\mu,\sigma)\ \sigma$$

$$\text{(b)} \quad m_2(\mu,\sigma^2) = \mu^2 + 2\mu Q(\mu,\sigma)\sigma + \sigma^2(1 + Q(\mu,\sigma)((w^r - \mu)/\sigma))$$

$$\text{(c)} \quad m_3(\mu,\sigma^2) = \mu^3 + Q(\mu, \sigma)\{ 3\sigma\mu(w^r - \mu) + 2\mu^2\sigma + 2\sigma^3 + (w^r - \mu)^2\sigma^2\} + 3\mu\sigma^2.$$

Gamma

Let offers be distributed Ga(α, β). The vector θ is then defined as $(\alpha, \beta)'$, α, $\beta > 0$. The density for W is

$$\text{(C.13)} \qquad f(w|\ \alpha,\ \beta) = \frac{(w/\beta)^{\alpha-1}[\exp(-w/\beta)]}{\beta\ \Gamma(\alpha)}\ ,\ w > 0,$$

where $\Gamma(\cdot)$ is the Gamma function defined as

$$\text{(C.14)} \qquad \Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt\ .$$

The density for accepted wages may be written as

$$\text{(C.15)} \quad f^T(w|w \geq w^r,\ \alpha,\ \beta) = \frac{w^{\alpha-1}[\exp(-w/\beta)]}{\int_{w^r}^\infty t^{\alpha-1}\ [\exp(-t/\beta)]dt}\ ,$$

for $w \geq w^r$, $w > 0$. Defining

$$\text{(C.16)} \qquad F(w|\ x,\ \beta) = \int_0^w f(u|\ x,\ \beta)du\ ,$$

the moments for the truncated distribution may be defined as

$$\text{(C.17)} \quad \text{(a)}\ m_1(\alpha,\ \beta) = \alpha\beta\ \frac{[1 - F(w^r|\ \alpha + 1,\ \beta)]}{[1 - F(w^r|\ \alpha,\ \beta)]}$$

$$\text{(b)}\ m_2(\alpha,\ \beta) = \alpha(\alpha + 1)\beta^2\ \frac{[1 - F(w^r|\ \alpha + 2,\ \beta)]}{[1 - F(w^r|\ \alpha,\ \beta)]}$$

$$\text{(c)}\ m_3(\alpha,\ \beta) = \alpha(\alpha + 2)(\alpha + 1)\beta^3\ \frac{[1 - F(w^r|\ \alpha + 3,\ \beta)]}{[1 - F(w^r|\ \alpha,\ \beta)]}$$

In addition to the normal and gamma distributions, I also experiment with the lognormal. Let Y = ln(W), the natural logarithm of an offer, be distributed $N(\mu_L, \sigma_L^2)$. Then the vector Θ is $(\mu_L, \sigma_L^2)'$, $-\infty < \mu_L < \infty$ and $\sigma_L^2 > 0$. The density for log offers Y, the density for accepted log wages, and the moments of the truncated distribution $m(\mu_L, \sigma_L^2)$ (i.e., $E(Y^j|w \geq w^r)$) are defined as above for the normal case, with $y = \ln(w)$, $y^r = \ln(w^r)$ substituted for w^r, and $(\mu_L, \sigma_L^2)'$ substituted for $(\mu, \sigma^2)'$.

Appendix D

Data

SIPP has two objectives (1) to provide more comprehensive information on the economic situation of households and persons in the United States than available elsewhere and (2) to do this in a way that allows analysis of changes over time. While not flawless, the data available at this time from SIPP provide an incredibly rich source that goes quite far in the direction of fulfilling both goals. For empirical labor economics, SIPP surpasses alternative sources of U.S. micro data -- both cross-section and panel -- in many ways. The major weakness of SIPP is the difficulty (and computer time) involved in exploiting the relative wealth of information it offers. Both the survey design and available data files have a relatively complex structure and extracting desired information is consequently less straightforward than when working with alternative sources. This is particularly true when exploiting its longitudinal features.

In Section D.1, I provide an overview of the basic structure of SIPP and its major advantages and disadvantages when judged against alternative data sources often used in labor market analysis. Section D.2 turns to the specific SIPP data and procedure I use to construct my data set. The detail of this discussion may appear cumbersome, but even a quick reading of the section will reveal that it is warranted by the complexity of the SIPP design and instruments, the structure of the data files currently available to the public, and the many decisions consequently required to obtain the duration and accepted wage measurements desired here. The strengths and weaknesses of SIPP as a basis for dynamic labor market analysis, in particular, are a focus throughout these first two sections.

Section D.1. The Basic Structure of the SIPP[1]

There are three basic components to the SIPP sample design: the Panel, the Rotation Group, and the Wave. Understanding these is fundamental to understanding SIPP.

A SIPP Panel is a multi-stage stratified random sample of the noninstitutionalized resident population of the U.S.[2] The household serves as the designated "ultimate sampling unit" and all persons of age fifteen or more residing in such designated units at the initial interview are eligible for inclusion in the Panel. The design calls for all persons in the Panel to be followed for the duration of the survey except for periods of institutionalization or residency outside the U.S.[3] Persons who become members of the designated households over the sampling period are added to the sample and followed thereafter, as are members of households formed by movers over the

---

[1]This section is intended to serve only as an overview of the SIPP. It is provided in light of the relatively recent release of the SIPP data and consequent lack of familiarity with the survey anticipated on the part of the reader. The general information contained here has been collected from a number of Bureau of the Census publications, including Nelson, et. al. (1985), Bureau of the Census (1985), and the Technical Documentation for Waves 1 to 4 for the 1984 Panel. The reader interested in using SIPP is advised to start with the first of these and David (1985), a separate source of useful information on SIPP (including discussions its potential for use in specific fields). In particular, see Fields and Jakubson (1985) in this volume for a preliminary discussion of SIPP's potential for labor economics.

[2] This restriction on the SIPP universe means that persons living in group quarters such as dormitories and convents are thus included in the design, while those living in military barracks, nursing homes, and correctional facilities are excluded. A detailed description of the multi-stage sampling plan may be found in the SIPP User's Guide (Bureau of the Census, 1985).

[3]The design initially stipulated that the survey cover a period of at least 2.5 years, although this has been altered for some households as discussed below.

age of fifteen. Children under the age of fifteen remain in the sample as long as they continue to reside with sample adults.

A Rotation Group is a subsample of a Panel obtained by simply partitioning the Panel into four groups having nearly equal numbers of sampling units. The purpose of this structure is to smooth out the interviewing and data processing procedures. Each month over the sampling period for a Panel, a different Rotation Group is interviewed and the four months preceding a particular interview month serve as the reference period (i.e., the period to which basic interview questions pertain).

A Wave is a set of interviews that use the same survey instrument (i.e., questionnaire). A set of Core questions is repeated in each interview and different sets of questions that pertain to a variety of special topics are also asked in some Waves. In the SIPP terminology, the latter are referred to as Topical Modules and Education and Work History, Health and Disability, and Assets and Liabilities are among the many topics covered. The Core questionnaire itself consists of four sections entitled Labor Force and Recipiency, Earnings and Employment, Amounts, and Program Questions. These titles are self-explanatory insofar as the general subject matter of each section is concerned and discussion in the next section covers the structure of the first two in some detail.

The SIPP design calls for more than one Panel. Each is selected and treated independently, including some variation in the content of the survey instruments across Panels (both Core and Topical sections). The design for the 1984 Panel (the initial sample in the SIPP and that with which I work) consisted of approximately 26,000 households. Of these, approximately 21,000 were found occupied and thus eligible for a first interview in the Fall of 1983 when interviewing began for this Panel. Originally, the plan was to follow all persons in the first three Rotation Groups of the 1984 Panel for nine Waves and those in the fourth Rotation Group for eight.[4] In Wave 5, however, approximately 850 households were deleted from each of the Rotation Groups in response to budgetary cutbacks. Together with attrition, this left a sample of approximately 15,600 households to be followed through the remaining Waves.[5] Interviewing for a second Panel began in February 1985. Here, the designated sample size was 17,800 and 13,300 initial interviews took place. The designated sample size for subsequent Panels added in each February thereafter is the same. As for length, the SIPP design calls for eight Waves for three Rotation Groups and seven Waves for the fourth for the 1985 Panel and each of the subsequent Panels.

From even this brief overview of the structure of the SIPP design, a number of appealing aspects are obvious. Because of the overlap in the calendar sampling periods across Panels, cross-section samples that are about half the size of the monthly Current Population Survey (CPS) samples are available, while the SIPP offers much more detailed income and labor market activity information than even the March CPS (the survey in which retrospective labor market data are collected for the preceding year).[6] As for alternative panel data sets, even the

---

[4]The difference in the number of Waves across Rotation Groups is due to the desire to have the Wave 6 interviews take place in the months of May through August since this Wave includes a Topical Module on Taxes. Rotation Group 4 is not interviewed in Wave 2.

[5]Note that this translates into a smaller sample to be potentially followed for the full duration of the Panel's sampling period since the some of the households had been added over the first four Waves.

[6]The designated sample size for the monthly CPS is approximately 71,000 household units and about 58,000 are typically interviewed. There is some degree of longitudinality in the CPS data in that household addresses (regardless of occupant changes) are interviewed for four months, rotated out for eight months, and then interviewed for four more. The fact that the housing unit and not the occupants is followed is obviously undesirable if longitudinal data are desired. See Ryscavage and Bregger (1985) for a detailed comparison of the SIPP and CPS (with a particular focus on

smaller SIPP Panel sizes are much larger than the sample sizes in the Panel Survey of Income Dynamics (PSID), the National Longitudinal Survey (NLS), the Employment Opportunity Pilot Project (EOPP), and Income Maintenance Experiment (SIME-DIME) data sets, for example. Also, the SIPP universe is not restricted to a particular segment of the population (other than the relatively loose restriction to noninstitutionalized U.S. residents), a feature shared with the CPS but none of these other surveys.[7]

Other features of the SIPP design that distinguish it from alternative panel surveys are the shorter reference period used at each interview and the detail and checking procedures used in the measurement of both labor and nonlabor income and labor force activity. (The latter, for example, is measured for each week in the reference period.) There are also features of the SIPP data that follow simply from its operating procedures. In the SIPP interviewing process, telephone interviews are used only in about five percent of all household interviews and proxy respondents are typically used for just over a third of all Panel members age fifteen or more. Both of these rates are relatively low and they affect the quantity of data collected, as well as its quality. For example, the item nonresponse rate for the wage and salary questions for Wave 1 of the 1984 Panel is 6.5 percent, with the rate for self-respondents being 4.6 percent versus a rate of 9.0 percent for proxies. As in all of the alternative surveys mentioned, SIPP interviews can collect only what respondents recall or wish to

---

differences in labor force concepts across the two).

[7]The NLS is structured so that particular age-sex cohorts are followed over time and there is an oversampling of the low-income population, as there is in the PSID, EOPP, and SIME-DIME samples. In the PSID, it is possible to identify the random subsample within the full sample, but restricting the sample in this way decreases the sample size significantly.

report. Social Security numbers are collected in the SIPP (and checked) and there are extensive plans to use these to match SIPP interview data with other data sources such as governmental administrative records and establishment records. The latter plan is of particular importance for empirical labor economics since the SIPP design (like the alternative surveys cited above) is restricted to households and thus to the supply side of the labor market.[8]

Despite all of these very positive attributes, the SIPP is nevertheless not perfect. In terms of disadvantages of its design and survey instruments, perhaps the most important is the length of the sampling period. Although households are followed over time, the length of the sampling period is relatively short. In particular, it is much shorter than the NLS, the PSID, and the SIME-DIME sampling periods.[9] Making matters worse, workers have no "history" upon entering the survey. Some information on a person's past may be obtained from Topical Modules, but only a limited amount. Furthermore, information collected in this way for the 1984 Panel is not collected until Wave 3 interviews or beyond and it is retrospective from the time that the questions are asked.[10]

Beyond these general problems with the SIPP design, there is the practical matter of the complexity of the data set that follows from the enormous quantity

---

[8]See Sater (1985) for some preliminary work on the matching of SIPP with supplementary procedures.

[9]The PSID has followed households since 1967 to the present and the NLS, which has a demographic cohort sample structure, has followed cohorts of individuals for up to 16 years. The SIME-DIME data follow households for a period of forty-eight months.

[10]The Work History module, for example, is in Wave 3 for the 1984 Panel. There is no work history topical module for the 1985 Panel, which is unfortunate. On the other hand, the work history topical module appears in Wave 2 for the 1986 Panel and it collects data retrospectively from the start of the survey.

of information collected for each individual. Multiple measurements are reported for many very similar but nonidentical items within Waves. Extreme care must therefore be exercised by the SIPP data user in order to avoid measuring something totally different from that intended. In exploiting the longitudinal features of the SIPP, the structure of the Public Use Files currently available compounds this. The data are available for all Waves for the 1984 Panel, but only in the form of cross-section files for each Wave. Longitudinal Research Files have been prepared that cover the first three Waves for all persons, but there is no plan to release files that contain all data for individuals over the entire sampling period.[11] The best that the Census Bureau plans to make directly available for the entire period are data aggregated to the monthly level of labor force activity.

The significance of these practical aspects of the SIPP--insofar as they may affect efforts to exploit even part of its labor market activity and earnings data--are addressed in greater detail in the next section.

Section D.2. Extracting Duration and Offer Data from SIPP

The data with which I work are taken from the Rectangular Public Use Files for Waves 1 to 4 for the 1984 Panel.[12] These data cover a period of up to

---

[11]A "Research File" is to be distinguished from a Public Use File in that the former is not an official Census Bureau data set while the latter is official. See Coder, et. al. (1987) for preliminary findings and a description of the first Longitudinal Research File. This File does not contain weekly labor force activity data, although an Extended Version is currently being prepared which will contain the weekly data by month but only for the first three Waves.

[12]In the Rectangular Files, each record is for a person. The data are also available in a Relational format where each record pertains to a household.

sixteen months for individuals in the first three Rotation Groups of the Panel and up to twelve months for those in the fourth. Figure D.1 displays the interview months and reference periods for each of the four Rotation Groups for this part of the survey period.[13]

In approaching these data, my primary interest was in obtaining measurements on: (i) basic demographic characteristics (age, race, sex, education, marital status); (ii) the length of the first initialized spell of joblessness observed (either uncensored or right censored); (iii) the characteristics of offers accepted at the end of uncensored first spells of joblessness (i.e., hourly earnings, occupation, industry, and usual hours); (iv) the length of an initialized spell of employment; and (v) the amount of weekly state Unemployment Insurance benefits received during the spell of joblessness. The first three were viewed as most important for my purposes.

While most of the requisite information is provided in SIPP, the structure and size of the Public Use files and content of the survey instruments preclude direct measurement of any of these variables (even demographic characteristics). A fairly complicated three stage procedure was devised to obtain measurements on all. First, each of the Wave files is read separately and raw data collected. The data for each Wave is then processed separately in order to associate accepted offers with the appropriate transitions. Finally, observations for each individual across Waves are merged using the SIPP identification system and, thereafter, variables remeasured.

---

[13]These vary across groups because of the variation in numbers of weeks across reference calendar months. The SIPP files include the number of weeks in each reference month for each individual in each Wave file.

but not all weeks.[19] (See ANYLOOK and ALLOOK.) After this partitioning, SIPP identification information and data on demographic characteristics at the time of the interview, the number of weeks in each reference month, State Unemployment Insurance (UI) recipiency status for the entire period, and UI benefit amounts received in each reference month are collected for all nontransition-nonemployed persons.[20]

SIPP identification information, weeks per reference month, and data on demographic characteristics at the time of the interview are also collected for workers who had a job in the reference period. Beyond this, data collected depends on whether or not a person experienced a spell of joblessness or not, i.e., experienced a transition. (See ALLJOB.) For those in the nontransition-employed group, information pertaining to the job held during the first week of the reference period is collected. (Details on available job characteristic data are provided in the Stage 2 discussion.) For those who experienced a transition, data is first collected on week by week job status and checked to determine whether or not a transition into a job occurred. (See JHIST.[21]) If a worker experienced only one transition and this was out of a job, then both data for the job initially held and UI data are collected. Otherwise, all available job and UI data are collected.

---

[19]The purpose of sorting the nontransition jobless group at this stage was primarily geared toward being able to determine whether or not a person had participated in any Wave reference period when merging the separate Waves' data. The label Discouraged should be interpreted loosely here (i.e., not as being in full agreement with the conventional definition of the term). New entrants and reentrants are also included here under this heading.

[20]Monthly amounts are the least aggregated measure available for UI benefits and almost all other income amounts. The latter is discussed below.

[21]It should be noted that having a job takes precedence over looking for a job in these weekly status observations.

My first sorting is therefore based on whether a person had a job in any weeks. (The actual question that generated the data is la in Table D.1 which I have labeled ANYJOB. Figure D.2 displays the partitioning scheme at this point and beyond using labels assigned to the questions in Table D.1).

Looking at those who were without a job for the entire reference period first, I next consider whether or not individuals participated in the labor market during any weeks. Conventional CPS definitions are invoked for unemployment and nonparticipation status; discouraged worker status is assigned for the Wave if labor market[18] participation involved looking for a job in some

---

regarded as a first round use of the SIPP data for empirical work in the search context. Overall, there are twenty-five questions pertaining to labor market status (many are asked for each week) and it is potentially possible to determine whether, for each week: a person had a job and worked, had a job but was absent without pay and why, didn't have a job but looked, looked but wouldn't take a job and why, didn't look but would have taken a job if offered, or did not participate at all. Determining status at this degree of accuracy involves merging responses to all questions for each week for each individual. At this time, it is planned as a future research project.

[18]The uppercase names are assigned by me. The question numbers are those that appear in the questionnaire. The numbers following these in parenthesis are coding numbers and these also appear in the questionnaire. When a question has been coded precisely as it appears on the questionnaire, this number allows one to determine what question is used and the context.

Aside from the remarks in brackets following the possible responses, the survey questions appear here exactly as they appear in Section 1 (Labor Force and Recipiency) of the 1984 Panel Wave 1-4 Questionnaires. The numbers in parenthesis are given in the Questionnaire and Codebook when data are reported as collected. This is quite useful to know since (1) the information collected often appears in an edited form (e.g., aggregated over time) on the Public Use Files, (2) the context of a question may influence a response, and (3) auxiliary questions and information may be used by the interviewer but not reported in the Codebook (e.g., the note below la). Working with the Codebook, Interviewer Instructions, and Questionnaire simultaneously allows one to avoid at least some of the ambiguity potentially involved with data interpretation. (Consider question 4 given in Table 2.1., for example.) As for the overall potential for ambiguity with SIPP, a signal is given by the fact that there are twenty-five questions pertaining to labor market status in Section 1 of the Core Questionnaire and twenty-seven pages devoted to them in the Interviewer Instructions.

of data possible be collected for each individual.[16] Since the latter depends on labor market status and, more important, on observed transitions between labor market states during the reference period, a multi-level partition of the sample based on these factors is used and pertinent data collected for each group therein as follows.

First, the specification of the labor market state space with which I work is based on having a job versus not, where a "job" is defined in SIPP as follows:

> A job exists if there is a definite arrangement for regular work for pay... A formal, definite arrangement with one or more employers to work a specified number of hours a week or days a month but on an irregular schedule during the week or month is also considered a job...(Bureau of the Census, 1985, p. E3-4)

> Consider a person as having a job or business if he/she was at work or had a definite arrangement with an employer to return to work at some particular date in the future. Consider a person as having a job or business when he/she is only temporarily absent from work due to illness, bad weather, vacation, layoff, or for some other relatively short period of time when it is clear the employer is holding the job open expecting his/her return ....(U.S. Bureau of the Census, 1985, p. E3-14)[17]

---

[16]There are over 30,000 individuals in the larger Wave 1, 3 and 4 nonreject samples. Precise breakdowns are given in the summary section below.

17It is important to distinguish this categorization from "working" versus "not working." Definite arrangements to return to work after some specified amount of time are included under this job definition (e.g., following a summer vacation or a paid leave of absence to attend school or training), while an arrangement to be called to work when work is available is not (e.g., listing with a union hiring hall or nurses' register). As for layoffs, workers on temporary layoff are viewed as having a job, while those on permanent layoff are not. This is possible because of the distinction made between working and absence without pay when a job is held. A layoff is treated as temporary in SIPP if it occurs

> because of material shortages, lack of work, inventory taking, plant remodeling, installation of machinery, or other similar changes (U.S. Bureau of the Census, 1985, p. E3-16.)

Given the search model underlying my analysis and the SIPP "job" definition, looking at spells of joblessness seems an appropriate choice and I work primarily with responses to SIPP questions pertaining to this status. It is important to note, however, that the information I use does not fully exploit the labor market activity information provided, but should instead be

Since numerous decisions are required within each of these stages, I describe them in (perhaps boring, but nevertheless what seems warranted) detail below so that the reader will understand exactly what has been analyzed when viewing the estimation results reported in the text.

Stage 1: Reading Each Wave

The subsample of the 1984 Panel with which I work consists of all persons between the ages of 16 and 70 at the time of the Wave 1 interview who during all observed Wave reference periods: (i) were not self-employed, (ii) did not report a disability that affected the amount or type of work that could be done, and (iii) did not report working as an unpaid worker in a family business. [14] Satisfaction of each of these criteria can change over time and thus across Waves. Consequently, the first step in reading each Wave file is to identify "rejects" and record their SIPP identification information for use in the merging stage.[15]

As for the remainder of a Wave sample (i.e., potential members of my final sample), sample sizes effectively impose the requirement that the minimum amount

[14]In the analysis reported below, additional restrictions are imposed (e.g., the sample is restricted to persons who did not report being employed in either an agricultural occupation or industry at any time). These restrictions are discussed in Section 3.

[15]Individuals who either entered the Panel after the first Wave or left prior to the fourth are included in the final sample if the above criteria are satisfied for all observed Wave reference periods. The age criteria is checked using a year of birth restriction, i.e., beyond age, the reported year of birth must be between 1913 and 1967. Details on the SIPP identification system are presented in the discussion of Stage 3 below.

Stage 2: Measurement and Identification Within Waves

Once the necessary data have been selected for an individual, the next step is processing it. The content of the Core section of the SIPP questionnaire is such that direct measurement of spell durations in weeks within Waves is fairly straightforward (at least for the job and joblessness spells measured here). For the nontransition groups, the length of the reference period (measured as the cumulative number of weeks over the reference months) provides a left and right censored spell length. For workers in the transition group, the first uninitialized spell, the last right censored, and (when observed) the first completed spells of having a job and joblessness in between can all be measured using the week by week observations on job status (JHIST).[22]

Once durations have been measured, uninitialized and first initialized spells of joblessness completed within the Wave reference period are linked with the appropriate accepted offers (i.e., earnings, hours, occupation, and industry for the job entered). When an individual starts the reference period with a job, the characteristics for that job must also be linked with it. Once a job is labelled in either way, a measure of average hourly earnings is then calculated when hourly earnings are not observed directly. The second of these steps - measuring wages or, to be more precise, average hourly earnings for employment with different employers discussed during the interview - is quite simple, but the structure and content of the questionnaire makes linking

[22]Employment date data (described below) is used as a supplement in some cases for first round measurements for practical reasons, but the duration measurements in the end are checked for consistency with the weekly status data.

accepted offers with completed spells of joblessness less than a straightforward exercise for large numbers of workers in each of the Wave samples. In some cases, it is impossible.

Each person who responded that he or she had a job during the reference period (i.e., ANYJOB=1) was then asked a series of questions about up to two places of employment in Section 2 of the Wave Questionnaire (Earnings and Employment). To start, the worker was asked if he or she had worked for an employer, been self-employed or both.[23] If the worker responded that he or she had (as all nontransition-employed and transition workers in my sample did), then the worker was asked the number of employers (up to three). Questions followed that pertained to the occupation, industry, hours, whether or not the job paid by the hour, hourly earnings if it did, monthly earnings, and the calendar period of employment when employed by a first employer.[24] When relevant, these questions were repeated for a second employer. Table D.2

[23]This is the question used to sort out rejects. Even if the worker reported that he or she had both worked for an employer and been self-employed, the worker is regarded as a reject here.

[24]The reader may wonder why this employer calendar information is not used as the primary source of duration data and the weekly observations as the secondary source, since this would allow measurement in days as opposed to weeks. First, these employment data are not problem free, as I discuss below. Second, as noted earlier, this is a first round attempt to exploit the longitudinal features of SIPP. Exploring possibilities such as this is an area for future research.

reproduces the actual questions that generated the data with which I work.[25]

The separate question on hourly earnings (6) represents an extremely valuable source of data. A separate question such as this is not typically provided in surveys and responses to it are available for a large proportion of each of the Wave samples. One problem with the data generated by this question is that a response may include raises received over the term of employment. Because of the short reference periods, this should not be a serious issue. As for workers who were not paid by the hour, the responses to the "usual" hours and monthly earnings questions (4 and 8), together with weeks employed by each employer in each reference month (calculated by the Bureau using responses to the employment period question (3)) provide fairly good data for calculating average hourly earnings. Obvious sources of measurement error in making such calculations are (i) the use of "usual" hours for all weeks, (ii) the failure to distinguish between base pay and overtime or bonus pay in reported monthly pay, and (iii) the lack of information with respect to lags in the actual receipt of earnings. These problems are no more serious than measurement error problems encountered with alternative data sets. They might even be described as less serious than usual because of the availability of monthly earnings (versus annual) and the shorter reference period over which such information

[25]Aside from occupation and industry questions and a question regarding frequency of labor income receipt, questions 2, 3, 4, 6, and 8 actually represent all questions asked about the first job and they were simply repeated for the second job if there had been more than one employer.

It should be noted that in the coding process for data collected after Wave 1, job related information was imputed by the Bureau of the Census when not reported by the individual. Imputation flags are provided in the Public Use Files, however, and these were used to blank out all imputed data in constructing the data set here.

must be recalled by respondents (typically a year).[26]

As for linking spells and jobs, there is no problem when there is only one employer and only one transition. Also, when a person experiences two job spells with two different employers and these are the only employers (i.e., either an initial and an entered job or two entered jobs), there is no problem. The calendar information provided by question 3b is sufficient to sort everything out. The first reference month in which a job was held and thus the first month's earnings for that job may be determined. Accepted average hourly earnings for accepted offers may then be calculated. In the event that no earnings are reported for the first month (due to lags in the pay periods, for example), the second month's earnings may be used if the job is accepted in the first three months of the reference period and it is held into a second month.

For all remaining cases (which fortunately represent a minority), the situation is more complicated. The sources of the complications are (i) the maximum number of two employers being discussed and (ii) the nature of the two questions that determine which are discussed and in what order, questions 2a and 10a in Table D.2. The wording of the questions is such that chronological order and time spent working for different employers over the reference period simultaneously determine the choice of the jobs discussed and the order. If there are two or more employers, starting dates with each employer discussed must be checked against the starting dates of the job spells of interest. For persons with three or more employers, this is obviously necessary because of the

---

[26]The monthly CPS is an exception on both counts. Earnings data are collected for the week preceding the survey, but only for one fourth of the monthly sample (the outgoing Rotation Group). Additional problems with exploiting the longitudinal feature of the CPS arise because of the residence based sample design.

"or most recent" part of the selection question (1a). The desired information is not always reported. Beyond this, under the definition of a "job" invoked here, a change of employers is not required across job spells. For example, a person may experience an indefinite layoff from a primary place of employment and return to that employer.[27] While with that employer, either before or after the observed spell of joblessness, the worker may simultaneously hold a second job (i.e., "moonlight"). A second employer may therefore be discussed in the survey but not correspond to either of the job spells in the accepted offer sense that is of interest to me. In such situations, I treat the return to the primary employer as an accepted offer. Insofar as the linking is concerned, if the employment dates and transition times allow me to determine the employer to whom jobs entered after joblessness pertain, I link. If there is uncertainty with respect to which employer the individual ends a spell of joblessness with, however, I do not link an offer. Once all linking that can be done has been done, I then proceed as in the simpler cases (i.e., calculate average hourly earnings for the first month, etc.).

Finally, for each transition out of joblessness, demographic characteristics and State UI benefits (where observed) are also linked. Focusing first on the demographics, the SIPP offers very detailed information on these. Age and marital status, for example, are both recorded at the time of the interview and for each month in the reference period. However, because of the amount of data that must be carried through -- regardless of a worker's

---

[27]Persons on indefinite layoff are regarded as being jobless, while persons on temporary layoffs are not. In the latter case, the job continues to exist, but work is interrupted. In the former, the worker does not deem the return to work as certain and would be willing to accept another job if offered. Obviously, there is room for ambiguity. In linking spells and offers, as in measuring durations, I maintain the SIPP job definition.

labor market status -- only the interview date demographic data are collected when reading the individual Wave files. These are then linked with all transitions experienced in the Wave reference period. As for UI benefits, amounts are reported for each month in the reference period, but the number of weeks for which they are received is not. To calculate a weekly amount, I divide the sum of all benefits over the entire reference period by the sum of the numbers of weeks that an individual is either jobless or absent without pay (the latter being included to cover temporary layoffs) over all months for which benefit amounts are reported. As for linking benefits with spells, I use the same Wave amount for all spells.[28]

Stage 3: Linking Observations Across Waves

After each of the Wave files has been read and the selected data processed, the last steps in constructing the data set are linking observations for individuals from the different Waves and then merging the data for each individual.

For the purpose of linking observations, SIPP provides a fourteen digit identification number for each individual in each Wave file.[29] When this number

---

[28]More fully exploiting the nonemployment income information in SIPP (as well as detailed demographic and labor market activity information) is planned for future work with SIPP. The rough benefit measure described above is not used in the central portion of the analysis reported below (Appendix F).

As noted in the text, the 1984 SIPP Core questionnaire does not adequately cover school enrollment. Only enrollment beyond high school is recorded at each Wave interview. This was changed for subsequent Panels.

[29]The identification number that appears in the Public Use File actually consists of three parts labelled SUID, PPENTRY, and PPNUM in the Codebooks. These refer to the sampling unit, entry level address within the sampling unit, and person number within the household for the individual. For protection of privacy, the first is given in a scrambled version on the Public Use files.

remains the same for a person across files, linking observations across Waves is simply a matter of matching identification numbers.

Unfortunately, this is not the case for all persons in the Panel, although it is for the majority of cases. The reason for this is that the identification number that appears in a particular Wave file depends in part on the sampling unit that the person has resided in over the reference period for that particular Wave. If this unit changes across Waves, a new number is assigned to the person.[30] In looking at two consecutive Waves files, for example, when a person record in the second of the two files cannot be matched with a record in the first, the numbering system is such that it allows one to determine whether or not the person is actually a new member of the Panel or simply in a new sampling unit. This is as far as one may go with identification numbers, however. As for recourse, too many individuals are affected in this way to match on other characteristics.[31]

A decision had to be made with respect to how such cases should be handled. Here, observations before and after such changes are treated as independent, i.e., a person in this situation appears as both a "leaver" and a "late entrant" in my final data set.

As indicated above, the ultimate objective in all of this is to obtain

---

[30]In a few cases (literally), individuals have been "lost" for a Wave because the date of the transition is too early to be in the old sampling unit and too late for the new one to be interviewed.

[31] When looking at the records for the first and second interviews for an individual, demographic characteristics such as sex and race may appear to change. While the former is actually possible, the latter is not given the current state of technology. What is actually going on here is a recode when incorrect information was recorded in the first interview. In merging observations for an individual, the second observations are given precedence over the first observations here for all characteristics that are time invariant (or at least treated as such in all empirical work).

observations on the length of first completed or initialized right censored spells of joblessness, the characteristics of offers accepted at the end of completed spells, and individual characteristics at the time of the transition. This objective together with practical considerations related to the sample sizes, individual Wave file sizes, and the potential for mismatches described above led me to link observations for individuals sequentially, accumulating information only until the desired amount of information is obtained.

More precisely, I start by linking observations for the first two Waves using the SIPP identification information. Once linked, I immediately implement a multilevel partition based on the type and amount of data I need to retain in moving into the next merge.

Since the rejection criteria pertain to the entire four Wave sampling period with which I work, the first level sort is simply reject elimination on the basis of both Waves, with identification data being retained for the rejects for future reference. The next level of sorting among the actual sample members is on the basis of whether of not a person appears in both Wave files.

Among those who do not appear in both, sorting is done according to where they do not appear and why. If an individual appears only in Wave 1 but he or she is not in Rotation Group 4, then the data for Wave 1 are simply relabelled as coming from the merge and the individual is regarded as done for my purposes. As noted above, Rotation Group 4 persons are excluded from Wave 2 by design and therefore kept separate. Their Wave 1 data are relabelled as coming from the merge stage, but unless an uncensored spell of joblessness has been completed and they are classified as done, they are placed in the subsample that continues into the link with Wave 3 data. New-entrants to the Panel in Wave 2 are treated in precisely the same manner as persons in Rotation Group 4.

For those who do appear in both Waves, the primary task is merging the duration data. This is fairly straightforward for most individuals and involves linking a Wave 1 right censored spell and Wave 2 left censored spell of the same kind (i.e., job or joblessness). When this produces a spell of joblessness and an accepted offer has been linked with the transition out of joblessness in Wave 2, then this offer is linked with this merge result. If a completed spell of joblessness is observed in Wave 1 or the merge result is an uncensored joblessness spell, then the person is classified as done. Alternatively, if the merge result is an uninitialized joblessness spell or an uncensored job spell, but an initialized and completed spell of joblessness is observed in Wave 2, then the person is classified as done. It may be the case that the person experiences no transition in either Wave. In this case, the individual is placed in the subsample that continues into the next merge (with Wave 3) with an uninitialized right censored spell equal to the sum of the lengths of the first two Waves combined.

In some cases, the duration observations at the end of the first and start of the second Waves do not match, i.e., an unobserved transition occurs at the "seam." One might suspect that there is inaccurate information reported for one or both sides of the seam, but there is effectively no choice as to how the situation can be handled. The spell that ends Wave 1 is treated as a completed spell and the spell that starts Wave 2 is treated as initialized. If the first is a spell of joblessness, then the job held at the start of Wave 2 is taken to be an accepted offer. Once the status of these seam spells has been determined, then the linking of offers, relabelling, and sorting according to whether individuals are done or not is carried out as above.

The next round of merging follows essentially the same pattern. The three subsamples from the first round merge are combined temporarily and then linked with the Wave 3 file. Rejection criteria are checked across all three Waves and rejects eliminated thereafter. Nonrejects classified as done at the end of the first round merge are then set apart and the remaining sample is partitioned in the same manner as in the first round merge. The data are also processed as in the first round merge, except that the results from the first round are treated as the first observation here. The third merge (i.e., with Wave 4) is carried out using precisely the same approach.

Figure D.1 SAMPLE DESIGN

Interview Months and Reference Periods

SIPP 1984 PANEL

Wave	Rotation Group 1	Rotation Group 2	Rotation Group 3	Rotation Group 4
1	OCTOBER 1983	NOVEMBER 1983	DECEMBER 1983	JANUARY 1984
	June 83 July 83 August 83 September 83	July 83 August 83 September 83 October 83	August 83 September 83 October 83 November 83	September 83 October 83 November 83 December 83
2	FEBRUARY 1984	MARCH 1984	APRIL 1984	
	October 83 November 83 December 83 January 84	November 83 December 83 January 84 February 84	December 83 January 84 February 84 March 84	
3	JUNE 1984	JULY 1984	AUGUST 1984	MAY 1984
	February 84 March 84 April 84 May 84	March 84 April 84 May 84 June 84	April 84 May 84 June 84 July 84	January 84 February 84 March 84 April 84
4	OCTOBER 1984	NOVEMBER 1984	DECEMBER 1984	SEPTEMBER 1984
	June 84 July 84 August 84 September 84	July 84 August 84 September 84 October 84	August 84 September 84 October 84 November 84	May 84 June 84 July 84 August 84

Figure D.2 STAGE 1 PARTITION: LABOR MARKET ACTIVITY WITHIN A WAVE

```
                          ANYJOB
                            /\
                           /  \
                      No  /    \  Yes
                         /      \
                 ANYLOOK          ALLJOB
                   /\               /\
                  /  \             /  \
            No   /    \  Yes  No  /    \  Yes
                /      \         /      \
     Nonparticipant   ALLOOK  Transition   Employed
          (NP)          /\       (T)         (E)
                       /  \
                  No  /    \  Yes
                     /      \
            Discouraged    Unemployed
               (D)            (U)
```

T: Transition observed within Wave reference period

NP, D, U: Without Job for duration of Wave reference period

E: With Job for duration of Wave reference period

Table D.1 LABOR MARKET ACTIVITY QUESTIONS

ANYJOB 1.(SC1000) During the 4-month period outlined on this calendar, that is, from (4 months ago) thru (last month), did ... have a job or business, either full time or part time, even for only a few days?

1. Yes (Question 4) 2. No (Question 2a)

ANYLOOK 2a.(SC1002) Even though ... did not have a job during this period, did ... spend any time looking for work or on layoff from a job?

1. Yes (Question 2b) 2. No (Questions pertaining to reasons for nonparticipation)

2b. (SC1004) Please look at the calendar. In which weeks was ... looking for work or on layoff from a job?

17 or 18 fields pertaining to each week in the individual's reference period

ALLOOK All weeks in the reference period marked as looking or on layoff

Table D.1 - Continued

ALLJOB	4. (SC1056) Did ... have a job or business, either full or part time, during EACH of the weeks in this period? Note that the person did not have to work each week. 1. Yes (Calendar pertaining to full weeks absent without pay and job and earnings questions) 2. No (Question 6a)
JHIST	6a. (SC1100-34) Pleas look at the calendar. In which weeks did ... have a job or business? 17 or 18 fields pertaining to each week in the individual's reference period Calendar and questions pertaining to full weeks absent without pay and job and earnings questions

Table D.2 EARNINGS AND EMPLOYMENT QUESTIONS

1b. (SC1716) How many different employers did ... work for during the 4-month period?

1. 1 employer 2. 2 employers 3. 3 or more employers

2a. (SC2000) What is the name of the employer for whom ... worked during this 4-month period?

(If ... worked for more than one employer, enter the employer for whom ... worked the most hours during the 4-month period or the most recent employer.)

3a. (SC2014) Was ... employed by (Name of employer) during the entire 4-month period?

1. Yes [Question 4] 2. No

3b. (SC2016-2022) When was ... employed by (Name of employer) during this 4-month period?

FROM Month Day TO Month Day

4. (SC2024) How many hours per week did ... usually work at this job?

Hours [Topcoded at 99]

6. (SC2028) What was ...'s regular hourly pay rate at the end of (last month or "to" date in item 3b)?

[Topcoded at 99.99]

Table D.2 - Continued

8. (SC2032-2038) The next question is about the pay ... received from this job during the 4-month period. We need the most accurate figures you can provide. Be sure to include any tips, bonuses, overtime pay, or commissions.

   What was the total amount of pay that ... received BEFORE deductions on this job last month?

   2 months ago?

   3 months ago?

   4 months ago?

   [Topcoded so that the average over the 4 months does not imply an annual income above $100,000]

10a. (SC2100) What is the name of the other employer for whom ... worked during this 4-month period?

   (If ... worked for more than one employer, enter the employer for whom ... worked the second most hours during the 4-month period.)

Appendix E

Alternative Reservation Wage Estimates

Tables E.1-4 present results from estimation of the model using each of the first two order statistics as estimators for the reservation wage. The corresponding plots are presented in Figures E.1-8 and the distribution parameter estimates are given in Tables E.5-8.

The results are essentially the same as those reported for the average of the first two order statistics. That is, the levels of the acceptance probabilities appear sensitive to the distributional assumption invoked, but the results again provide evidence of a strong systematic relationship between the level of the transition rate and the level of the arrival rate under each specification for the offer distribution. The level of the acceptance probability does not appear to play an important role in producing observed variation in transition rates across groups.

Table E.1 GAMMA DISTRIBUTION: WR1 - FIRST ORDER STATISTIC

GROUP:	N WAGE	WR1 w_c^r	WAGE $E(w\|w \geq w_c^r)$	TAU τ_c	GPI $\pi_c(w_c^r)$	GDELTA δ_c	GCHI[a] χ^2
White Males:							
Ages 16-19	296	1.000	4.067	0.073	0.999	0.073	4.152
Ages 20-24	404	1.029	5.640	0.069	0.995	0.069	1.766
Ages 25-44	556	1.500	8.261	0.068	0.987	0.068	1.109
Ages 45-64	187	1.139	9.857	0.036	0.990	0.036	1.083
Nonwhite Males:							
Ages 16-19	44	2.179	3.667	0.059	0.904	0.065	5.197
Ages 20-24	66	1.089	4.445	0.064	0.998	0.064	3.657
Ages 25-44	89	1.189	6.849	0.048	0.991	0.048	2.836
Ages 45-64	25	2.269	8.739	0.020	0.974	0.020	1.744
White Females:							
Ages 16-19	240	1.000	3.792	0.067	0.999	0.068	1.038
Ages 20-24	363	1.059	4.493	0.063	0.990	0.064	1.480
Ages 25-44	598	1.000	5.842	0.051	0.986	0.052	5.529
Ages 45-64	226	1.329	6.043	0.033	0.984	0.033	7.607
Nonwhite Females:							
Ages 16-19	40	2.000	3.602	0.064	0.980	0.066	9.770
Ages 20-24	63	3.000	4.355	0.042	0.862	0.048	10.133
Ages 25-44	129	1.459	5.911	0.039	0.966	0.041	1.170
Ages 45-64	36	1.319	4.655	0.023	0.995	0.023	1.148

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

Table E.2 NORMAL DISTRIBUTION: WR1 - FIRST ORDER STATISTIC

GROUP:	N WAGE	WR1 w_c^r	WAGE $E(w \mid w \geq w_c^r)$	TAU τ_c	NPI $\pi_c(w_c^r)$	NDELTA δ_c	NCHI[a] χ^2
White Males:							
Ages 16-19	296	1.000	4.067	0.073	0.512	0.142	34.159
Ages 20-24	404	1.029	5.640	0.069	0.431	0.161	2.679
Ages 25-44	556	1.500	8.261	0.068	0.358	0.189	0.594
Ages 45-64	187	1.139	9.857	0.036	0.805	0.045	7.570
Nonwhite Males:							
Ages 16-19	44	2.179	3.667	0.059	0.297	0.199	0.992
Ages 20-24	66	1.089	4.445	0.064	0.450	0.142	0.755
Ages 25-44	89	1.189	6.849	0.048	0.442	0.109	2.112
Ages 45-64	25	2.269	8.739	0.020	0.448	0.044	0.172
White Females:							
Ages 16-19	240	1.000	3.792	0.067	0.422	0.160	47.521
Ages 20-24	363	1.059	4.493	0.063	0.274	0.231	0.080
Ages 25-44	598	1.000	5.842	0.051	0.370	0.138	0.146
Ages 45-64	226	1.329	6.043	0.033	0.403	0.082	109.497
Nonwhite Females:							
Ages 16-19	40	2.000	3.602	0.064	0.322	0.200	1.311
Ages 20-24	63	3.000	4.355	0.042	0.579	0.072	21.447
Ages 25-44	129	1.459	5.911	0.039	0.350	0.113	0.040
Ages 45-64	36	1.319	4.655	0.023	0.451	0.051	0.699

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

Table E.3 GAMMA DISTRIBUTION: WR2 - SECOND ORDER STATISTIC

GROUP:	N WAGE	WR2 w_c^r	WAGE $E(w \mid w \geq w_c^r)$	TAU τ_c	GPI $\pi_c(w_c^r)$	GDELTA δ_c	GCHI[a] χ^2
White Males:							
Ages 16-19	296	1.109	4.067	0.073	0.999	0.073	4.151
Ages 20-24	404	1.079	5.640	0.069	0.994	0.070	1.760
Ages 25-44	556	1.869	8.261	0.068	0.971	0.070	1.080
Ages 45-64	187	1.589	9.857	0.036	0.975	0.037	1.078
Nonwhite Males:							
Ages 16-19	44	2.319	3.667	0.059	0.938	0.063	1.711
Ages 20-24	66	1.250	4.445	0.064	0.995	0.064	3.631
Ages 25-44	89	2.079	6.849	0.048	0.881	0.054	1.519
Ages 45-64	25	2.659	8.739	0.020	0.962	0.020	2.235
White Females:							
Ages 16-19	240	1.109	3.792	0.067	0.998	0.068	1.038
Ages 20-24	363	1.069	4.493	0.063	0.990	0.064	1.479
Ages 25-44	598	1.059	5.842	0.051	0.982	0.052	5.506
Ages 45-64	226	1.750	6.043	0.033	0.939	0.035	6.741
Nonwhite Females:							
Ages 16-19	40	2.250	3.602	0.064	0.937	0.069	9.738
Ages 20-24	63	3.119	4.355	0.042	0.891	0.047	17.954
Ages 25-44	129	1.509	5.911	0.039	0.961	0.041	1.162
Ages 45-64	36	1.839	4.655	0.023	0.918	0.025	1.098

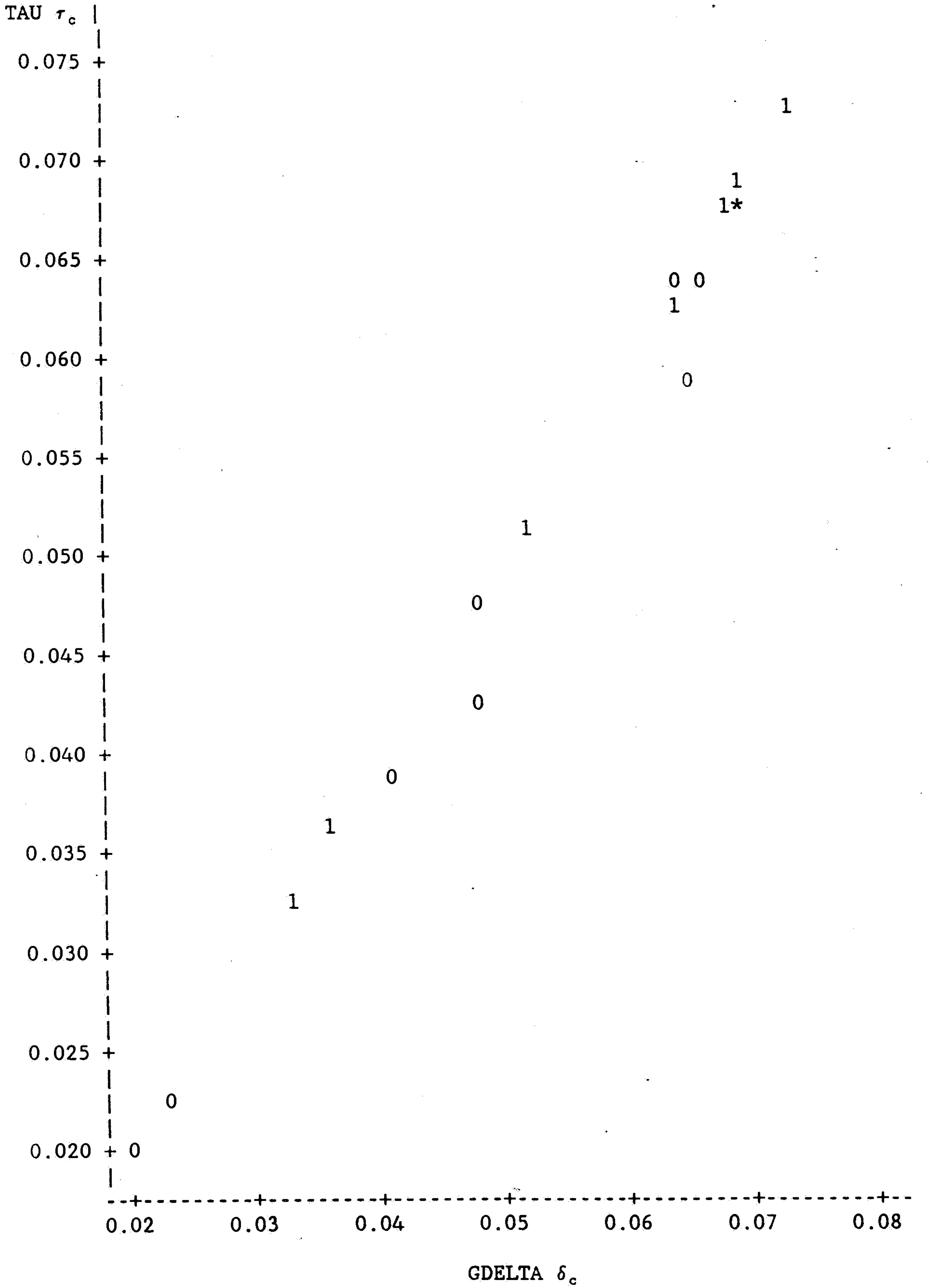
[a] The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

Table E.4 NORMAL DISTRIBUTION: WR2 - SECOND ORDER STATISTIC

GROUPS:	N WAGE	WR2 w_c^r	WAGE $E(w\|w \geq w_c^r)$	TAU τ_c	NPI $\pi_c(w_c^r)$	NDELTA δ_c	NCHI[a] χ^2
White Males:							
Ages 16-19	296	1.109	4.067	0.073	0.495	0.147	27.915
Ages 20-24	404	1.079	5.640	0.069	0.400	0.173	1.629
Ages 25-44	556	1.869	8.261	0.068	0.411	0.165	24.769
Ages 45-64	187	1.589	9.857	0.036	0.421	0.086	74.229
Nonwhite Males:							
Ages 16-19	44	2.319	3.667	0.059	0.246	0.240	0.496
Ages 20-24	66	1.250	4.445	0.064	0.423	0.151	0.430
Ages 25-44	89	2.079	6.849	0.048	0.340	0.141	0.108
Ages 45-64	25	2.659	8.739	0.020	0.411	0.048	0.049
White Females							
Ages 16-19	240	1.109	3.792	0.067	0.405	0.167	210.960
Ages 20-24	363	1.069	4.493	0.063	0.391	0.162	54.789
Ages 25-44	598	1.059	5.842	0.051	0.364	0.140	0.215
Ages 45-64	226	1.750	6.043	0.033	0.381	0.086	71.515
Nonwhite Females:							
Ages 16-19	40	2.250	3.602	0.064	0.242	0.267	0.453
Ages 20-24	63	3.119	4.355	0.042	0.200	0.209	14.116
Ages 25-44	129	1.509	5.911	0.039	0.345	0.115	0.055
Ages 45-64	36	1.839	4.655	0.023	0.342	0.067	0.257

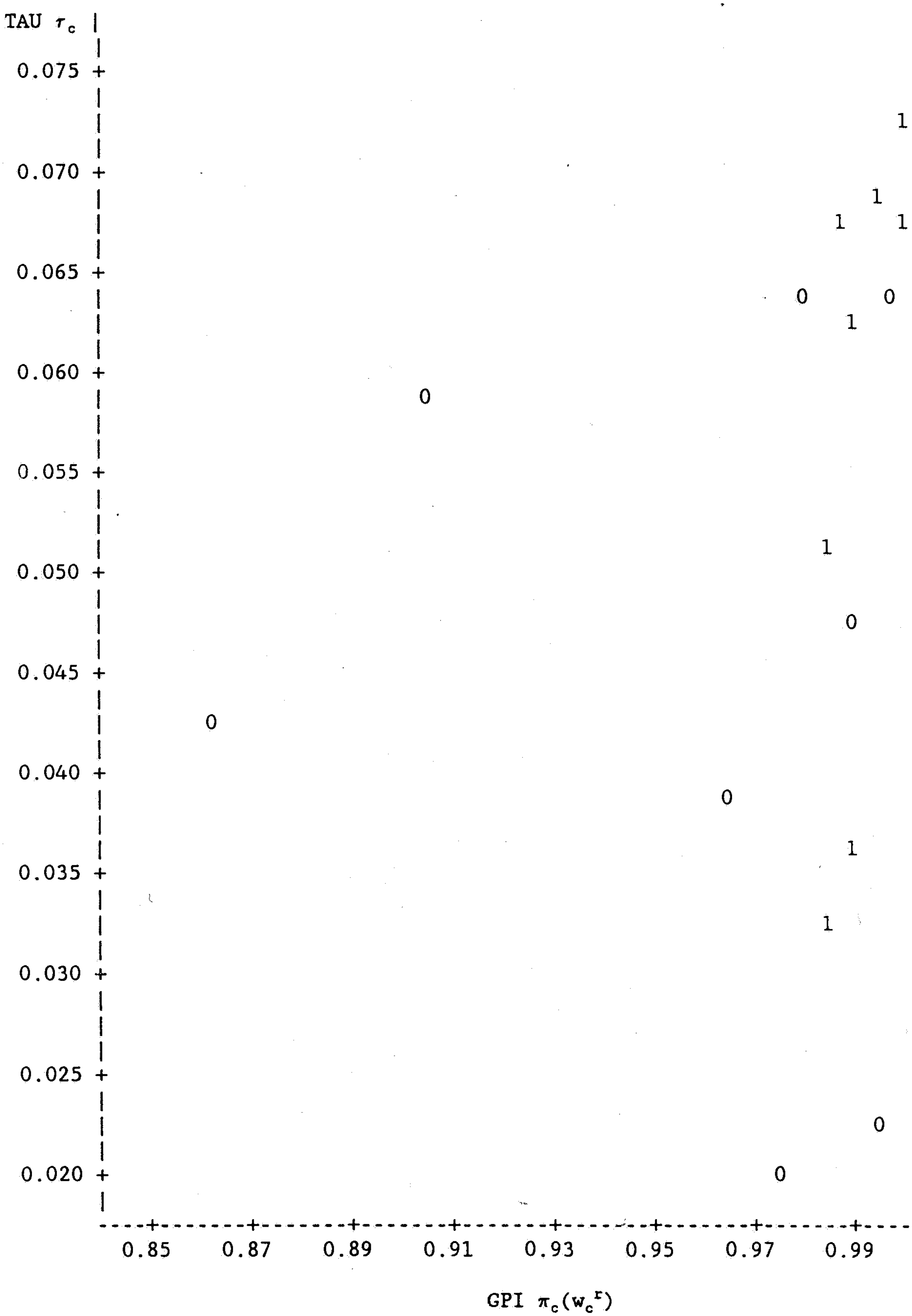
[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

Figure E.1 GAMMA DISTRIBUTION: WR1 - FIRST ORDER STATISTIC



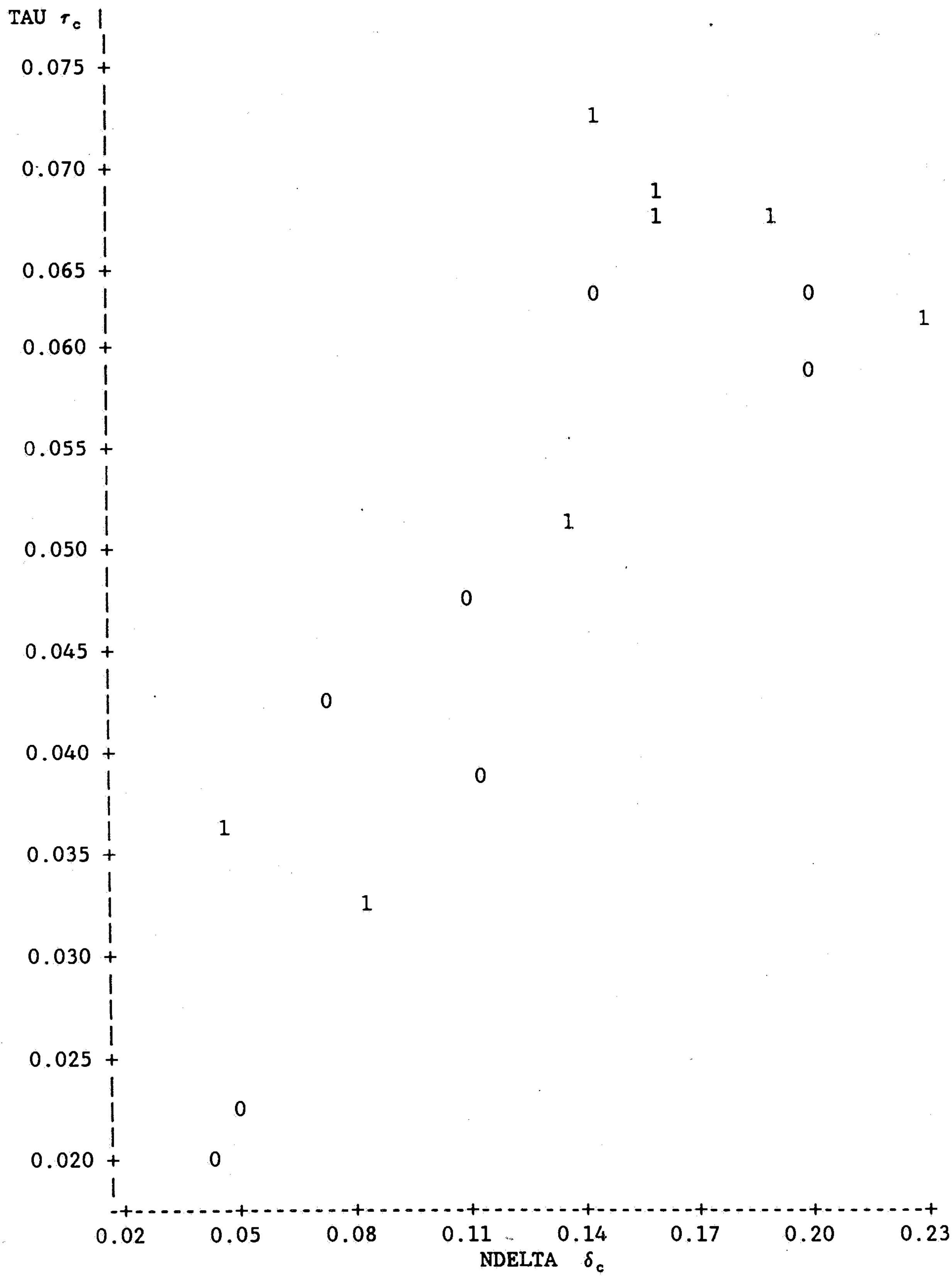
PLOT OF TAU*GDELTA SYMBOL IS RACE (White = 1) WITH * FOR TWO OBSERVATIONS

Figure E.2 GAMMA DISTRIBUTION: WR1 - FIRST ORDER STATISTIC



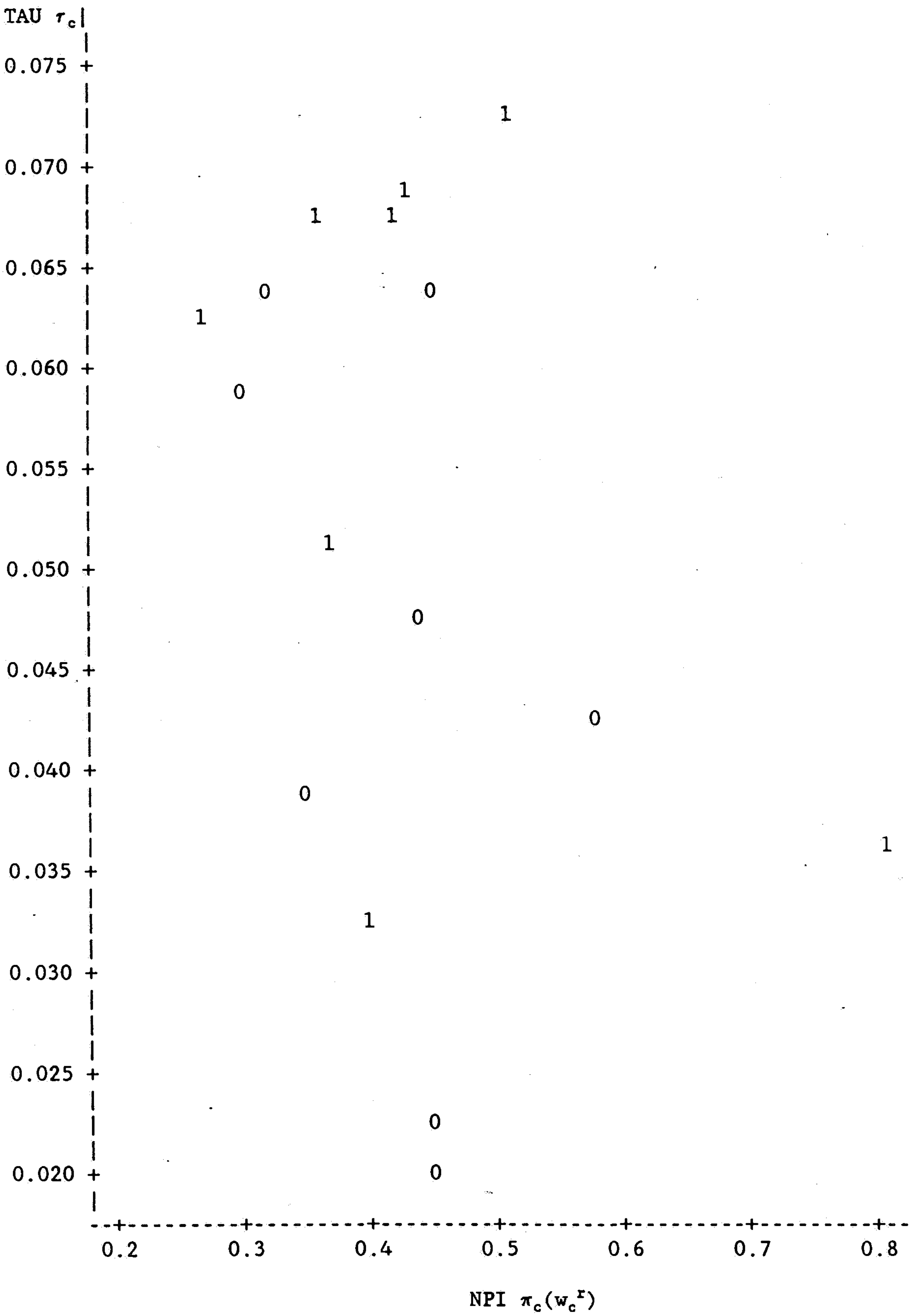
PLOT OF TAU*GPI SYMBOL IS RACE (White = 1)

Figure E.3 NORMAL DISTRIBUTION: WR1 - FIRST ORDER STATISTIC



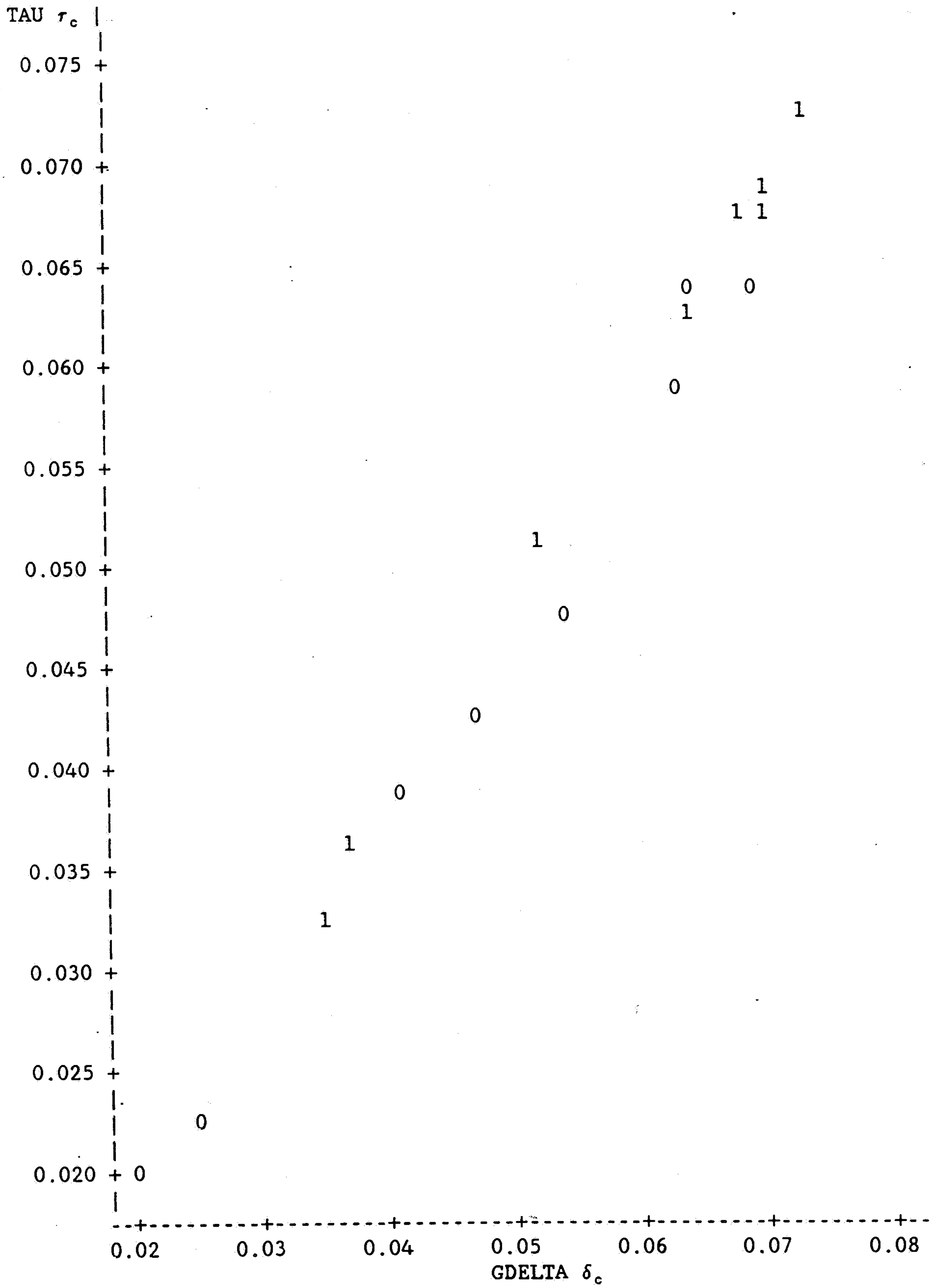
PLOT OF TAU*NDELTA SYMBOL IS RACE (White - 1)

Figure E.4 NORMAL DISTRIBUTION: WR1 – FIRST ORDER STATISTIC



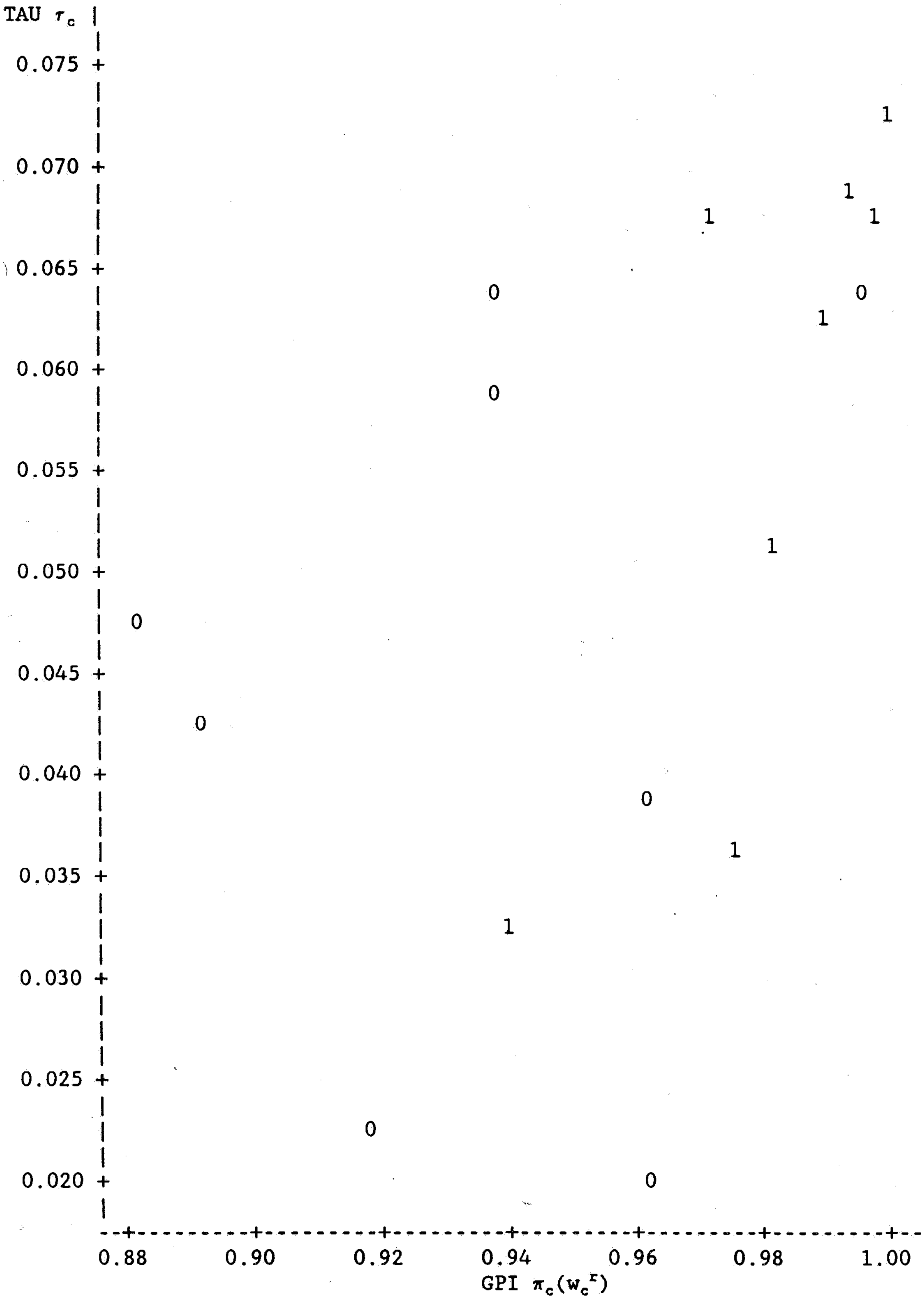
PLOT OF TAU*NPI SYMBOL IS RACE (White = 1)

Figure E.5 GAMMA DISTRIBUTION: WR2 - SECOND ORDER STATISTIC



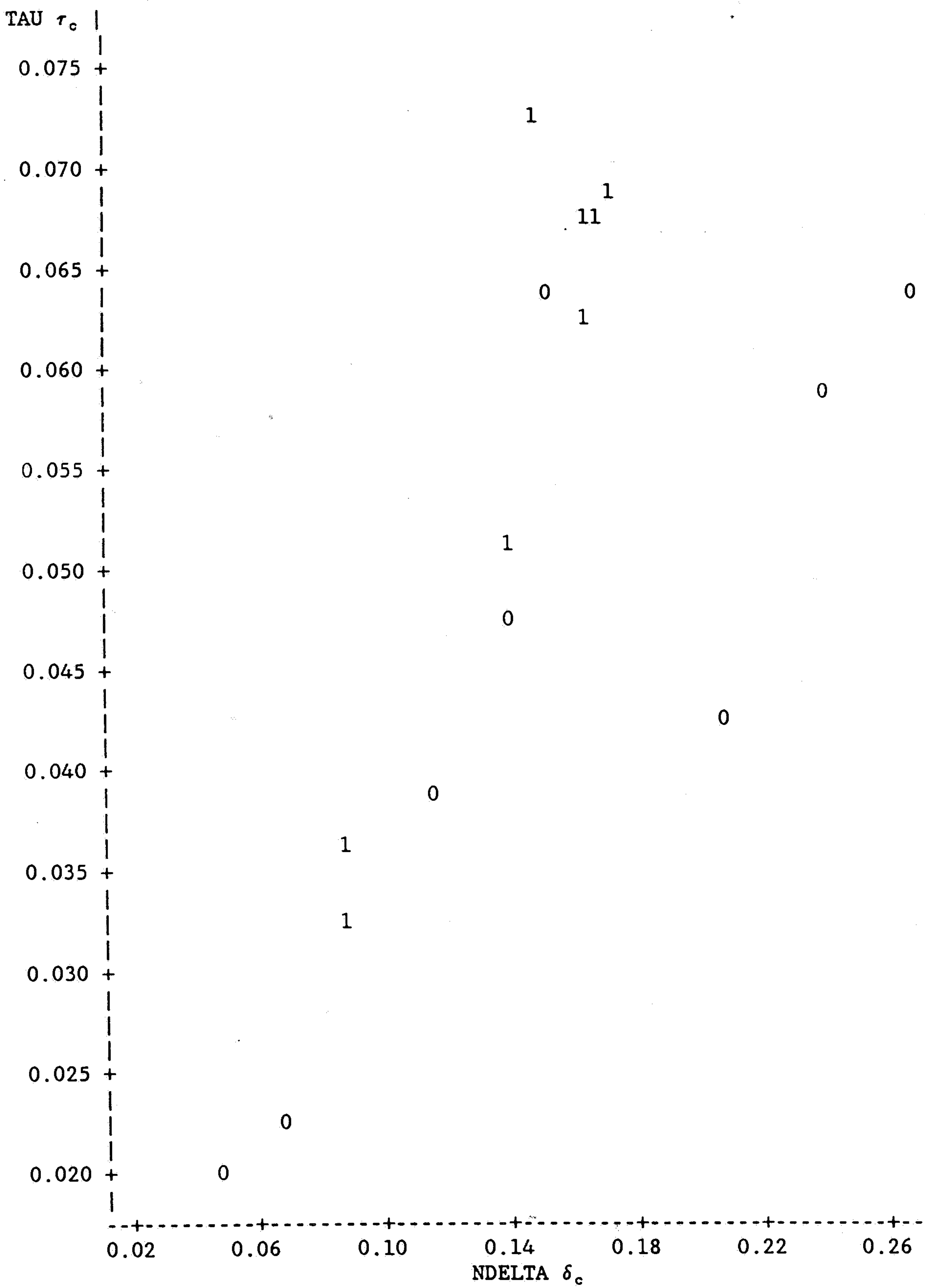
PLOT OF TAU*GDELTA SYMBOL IS RACE (WHITE=1)

Figure E.6 GAMMA DISTRIBUTION: WR2 = SECOND ORDER STATISTIC



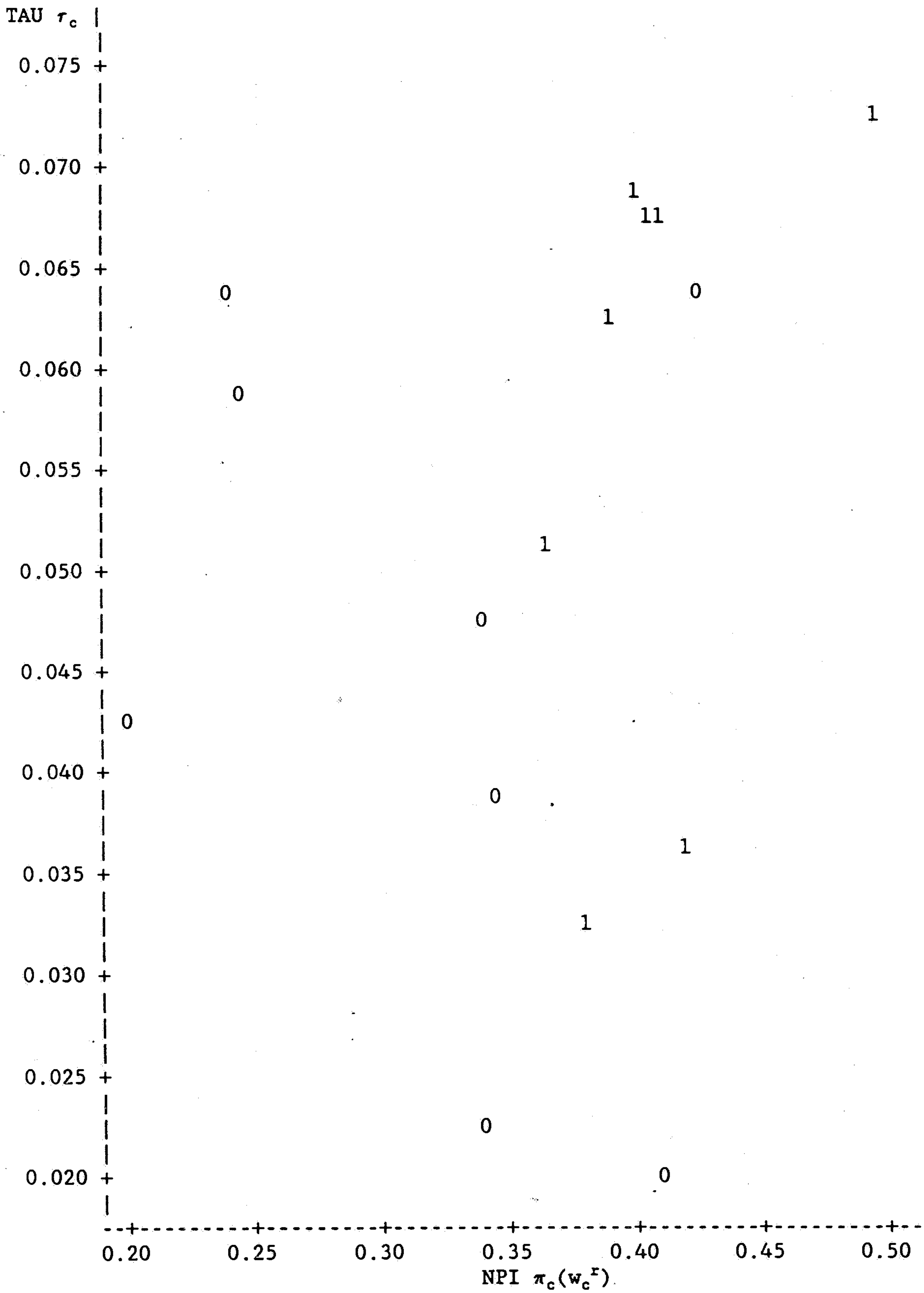
PLOT OF TAU*GPI SYMBOL IS RACE (WHITE=1)

Figure E.7 NORMAL DISTRIBUTION: WR2 - SECOND ORDER STATISTIC



PLOT OF TAU*NDELTA SYMBOL IS RACE (WHITE=1)

Figure E.8 NORMAL DISTRIBUTION: WR2 - SECOND ORDER STATISTIC



PLOT OF TAU*NPI SYMBOL IS RACE (WHITE=1)

Table E.5 GAMMA DISTRIBUTION: WR1 - FIRST ORDER STATISTIC PARAMETER ESTIMATES (Standard Errors in Parentheses)

GROUP:	N WAGE	ALPHA α_c	BETA β_c	GMU Ec(w)	GVAR Var(w)	GCHI[a] χ^2	GMAXF[b]
White Males:							
Ages 16-19	296	14.257	0.279	3.978	1.110	4.152	0.000
		(1.711)	(0.035)				
Ages 20-24	404	4.480	1.232	5.522	6.808	1.766	0.000
		(0.388)	(0.121)				
Ages 25-44	556	3.416	2.359	8.060	19.019	1.109	0.000
		(0.226)	(0.169)				
Ages 45-64	187	2.717	3.408	9.263	31.573	1.083	0.000
		(0.405)	(0.552)				
Nonwhite Males:							
Ages 16-19	44	10.074	0.350	3.527	1.235	5.197	0.110
		(0.429)	(0.016)				
Ages 20-24	66	8.112	0.501	4.065	2.037	3.657	0.000
		(2.133)	(0.138)				
Ages 25-44	89	3.808	1.684	6.414	10.804	2.836	0.000
		(0.661)	(0.333)				
Ages 45-64	25	3.934	2.127	8.370	17.804	1.744	0.000
		(1.446)	(0.719)				
White Females:							
Ages 16-19	240	9.123	0.401	3.563	1.471	1.038	0.000
		(1.601)	(0.072)				
Ages 20-24	363	4.885	0.886	4.330	3.837	1.480	0.000
		(0.872)	(0.164)				
Ages 25-44	598	3.237	1.709	5.534	9.459	5.529	0.000
		(0.272)	(0.158)				
Ages 45-64	226	4.029	1.383	5.576	7.717	7.607	0.000
		(0.485)	(0.184)				
Nonwhite Females:							
Ages 16-19	40	18.999	0.180	3.419	0.615	9.770	4.398
		(4.336)	(0.038)				
Ages 20-24	63	38.658	0.093	3.628	0.340	10.133	1.077
		(12.791)	(0.047)				
Ages 25-44	129	3.452	1.606	5.545	8.908	1.170	0.000
		(0.524)	(0.261)				
Ages 45-64	36	7.475	0.592	4.426	2.621	1.148	0.000
		(2.521)	(0.205)				

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

[b]MAX F is the maximum absolute value of the gradient at the reported estimates. In some cases, this value exceeds the desired value for convergence. In such cases, the sum of squares and reported parameter values were unchanging.

Table E.6 NORMAL DISTRIBUTION: WR1 - FIRST ORDER STATISTIC PARAMETER ESTIMATES (Standard Errors in Parenthesis)

GROUP:	N WAGE	NMU μ_c	NSIG σ_c	NCHI[a] χ^2	MAX F[b]
Nonwhite Females:					
Ages 16-19	40	0.773	2.668	1.311	0.007
		(0.149)	(0.014)		
Ages 20-24	63	3.329	1.648	21.447	4.096
		(0.020)	(0.055)		
Ages 25-44	129	-1.042	6.523	0.040	0.000
		(0.530)	(0.050)		
Ages 45-64	36	0.743	4.767	0.699	0.000
		(0.447)	(0.062)		
Nonwhite Males:					
Ages 16-19	44	0.830	2.539	0.992	0.009
		(0.107)	(0.003)		
Ages 20-24	66	0.501	4.742	0.755	0.000
		(0.702)	(0.225)		
Ages 25-44	89	0.065	7.735	2.112	0.385
		(4.938)	(2.502)		
Ages 45-64	25	1.091	9.051	0.172	0.000
		(0.580)	(0.184)		
White Females:					
Ages 16-19	240	0.207	4.072	47.521	0.658
		(1.656)	(0.773)		
Ages 20-24	363	-2.248	5.530	0.080	0.012
		(0.059)	(0.086)		
Ages 25-44	598	-1.307	6.953	0.146	0.000
		(0.154)	(0.053)		
Ages 45-64	226	-0.232	6.400	109.497	2.032
		(2.564)	(1.220)		
White Males:					
Ages 16-19	296	1.134	4.303	34.159	0.024
		(0.051)	(0.011)		
Ages 20-24	404	-0.103	6.575	2.679	0.002
		(1.820)	(0.906)		
Ages 25-44	556	-2.120	9.961	0.594	0.011
		(0.099)	(0.141)		
Ages 45-64	187	7.552	7.439	7.570	0.038
		(0.364)	(0.256)		

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

[b]MAX F is the maximum absolute value of the gradient at the reported estimates. In some cases, this value exceeds the desired value for convergence. In such cases, the sum of squares and reported parameter values were unchanging.

Table E.7 GAMMA DISTRIBUTION: WR2 = SECOND ORDER STATISTIC PARAMETER ESTIMATES (Standard Errors in Parentheses)

GROUP:	N WAGE	ALPHA α_c	BETA β_c	GMU $E_c(w)$	GVAR Var(w)	GCHI[a] χ^2	MAX F[b]
White Males:							
Ages 16-19	296	14.249	0.279	3.978	1.110	4.151	0.000
			(1.717)	(0.035)			
Ages 20-24	404	4.461	1.236	5.518	6.826	1.760	0.000
			(0.391)	(0.122)			
Ages 25-44	556	3.247	2.449	7.956	19.492	1.080	0.000
			(0.234)	(0.182)			
Ages 45-64	187	2.593	3.523	9.138	32.195	1.078	0.000
			(0.409)	(0.583)			
Nonwhite Males:							
Ages 16-19	44	17.121	0.206	3.530	0.727	1.711	1.011
			(1.005)	(0.014)			
Ages 20-24	66	7.865	0.516	4.063	2.099	3.631	0.000
			(2.390)	(0.161)			
Ages 25-44	89	2.375	2.591	6.155	15.950	1.519	0.003
			(0.510)	(0.559)			
Ages 45-64	25	4.087	2.066	8.448	17.460	2.235	0.000
			(1.596)	(0.724)			
White Females:							
Ages 16-19	240	9.071	0.403	3.660	1.477	1.038	0.000
			(1.634)	(0.074)			
Ages 20-24	363	4.878	0.887	4.329	3.841	1.479	0.000
			(0.875)	(0.165)			
Ages 25-44	598	3.203	1.723	5.520	9.513	5.506	0.000
			(0.274)	(0.161)			
Ages 45-64	226	3.326	1.637	5.447	8.922	6.741	0.000
			(0.492)	(0.246)			
Nonwhite Females:							
Ages 16-19	40	16.115	0.215	1.246	0.746	9.738	5.955
			(0.017)	(3.468)			
Ages 20-24	63	42.499	0.090	3.825	0.344	17.954	25.305
			(18.539)	(0.080)			
Ages 25-44	129	3.395	1.626	5.524	8.986	1.162	0.000
			(0.526)	(0.266)			
Ages 45-64	36	4.364	0.993	4.334	4.303	1.098	0.017
			(2.737)	(0.553)			

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

[b]MAX F is the maximum absolute value of the gradient at the reported estimates. In some cases, this value exceeds the desired value for convergence. In such cases, the sum of squares and reported parameter values were unchanging.

Table E.8 NORMAL DISTRIBUTION: WR2 - SECOND ORDER STATISTIC PARAMETER ESTIMATES (Standard Errors in Parenthesis)

GROUP:	N WAGE	NMU μ_c	NSIG σ_c	NCHI[a] χ^2	MAX F[b]
White Males:					
Ages 16-19	296	1.062	4.200	27.915	0.021
			(0.058)	(0.009)	
Ages 20-24	404	-0.573	6.594	1.629	0.019
			(1.506)	(0.538)	
Ages 25-44	556	-0.099	8.856	24.769	0.161
			(2.737)	(1.309)	
Ages 45-64	187	-0.331	9.677	74.229	0.232
			(11.574)	(5.193)	
Nonwhite Males:					
Ages 16-19	44	0.670	2.410	0.496	0.000
			(0.247)	(0.046)	
Ages 20-24	66	0.367	4.596	0.430	0.010
			(1.778)	(0.703)	
Ages 25-44	89	-0.786	6.965	0.108	0.000
			(0.863)	(0.136)	
Ages 45-64	25	0.724	8.678	0.049	0.000
			(1.391)	(0.191)	
White Females:					
Ages 16-19	240	0.240	3.645	210.960	1.911
			(1.969)	(0.880)	
Ages 20-24	363	-0.229	4.721	54.789	0.433
			(1.986)	(0.888)	
Ages 25-44	598	-1.318	6.889	0.215	0.000
			(0.153)	(0.053)	
Ages 45-64	226	-0.058	6.013	71.515	1.597
			(1.500)	(0.735)	
Nonwhite Females:					
Ages 16-19	40	0.550	2.429	0.453	0.000
			(1.087)	(0.317)	
Ages 20-24	63	1.192	2.298	14.116	0.453
			(0.147)	(0.028)	
Ages 25-44	129	-1.059	6.471	0.055	0.000
			(0.516)	(0.051)	
Ages 45-64	36	0.094	4.300	0.257	0.056
			(1.657)	(0.802)	

[a]The critical values for the chi-square statistic with one degree of freedom are 0.000982 and 5.02 at the 5 percent level and 0.0000393 and 7.88 at the 1 percent level.

[b]MAX F is the maximum absolute value of the gradient at the reported estimates. In some cases, this value exceeds the desired value for convergence. In such cases, the sum of squares and reported parameter values were unchanging.

Appendix F

Checks on the Reservation Wage Estimates

The optimality condition for the model set out in section 1 is given by

$$(F.1) \qquad (w_c^r - b_c)[r_c + a_c] = \{ (E_w[w|w \geq w_c^r] - w_c^r) \cdot ([1 - F_c(w_c^r)]\delta_c) \}.$$

As a check on how reasonable the estimates for the reservation wages are, on their own, I first calculate the net hourly unemployment income implied by the optimality condition for each of the reservation wage estimates. This is done using my estimates for the transition rates out of unemployment τ_c^*, the mean accepted wage distribution $E(w|w \geq w_c^r)^*$, and an estimate of the weekly discount rate. The discount rate is set equal to the sum of an estimate of the arrival rate of layoffs for each group based data on tenure in accepted jobs and an (albeit arbitrary) estimate for the rate of time preference $r_c^* + a_c^*$.[1] That is, I calculate

$$(F.2) \quad b_c^* = w_c^{r*} - \{[E(w|w \geq w_c^r)^* - w_c^{r*}]\ \tau_c^*\}/(r_c^* + a_c^*)$$

and check to see if the results look sensible. As another check, I move in the reverse direction and calculate the weekly discount rate implied by the optimality condition using an estimate for hourly unemployment income for each group, i.e., I calculate

$$(F.3) \quad (r_c + a_c)^* = \frac{[E(w|w > w_c^r)^* - w_c^{r*}] \cdot \tau_c^*}{w_c^{r*} - b_c^*}.$$

Table F.1 presents the results from these checks: (i) the discount rates implied under the optimality condition by the first and second order statistic

---

[1]My estimator for the transition rate out of employment is the maximum likelihood under an assumption of a constant transition rate between employment and unemployment (analogous to the estimator for τ_c, described in the text).

estimates for the reservation wage together with the r_c and $E[w|w \geq w_c^r]$ estimates and a rough measure of hourly UI benefits (DRATE1 and DRATE2), and (ii) the net hourly unemployment income levels implied the reservation wage estimates, the r_c and $E[w|w \geq w_c^r]$ estimates, and a rough estimate of the total discount rate equal to the sum of the transition rate out of employment TAUE and a weekly discount rate of 0.002 (IBEN1 and IBEN2).

Overall, these results are similar to those reported by Wolpin (1987), Flinn and Heckman (1982), and Stern (1989) in their analyses of young male workers. Therewith, they are somewhat sobering and call the results reported above into question. Specifically, the results suggest that the order statistic estimators provide estimates that are too low if the defining condition for the reservation wage is valid. Alternatively, they may be interpreted as implying that the model is misspecified, i.e., it is too simple for this demographic group partition of the sample. A third and more optimistic interpretation is that the benefit level and discount rate measures are too rough. There is no way of assigning a rate of time preference without using personal judgment. Obviously, positive net unemployment income estimates could be produced if the discount rate were set high enough. The benefit level estimates used here represent averages for those who receive benefits and, thus, they do not account for varying probabilities of receipt across groups. Also, as discussed in the previous section, the benefit estimates used here are measured only imprecisely.

In the less favorable light of the results from these reservation wage checks, I experimented in two directions. First, I experimented with estimation using an alternative estimator for the reservation wage. Specifically, for each group, I calculated the solution to the optimality condition for the reservation wage using the measures for the weekly discount rate and benefit levels used in the checks above. (These measures are given in Table F.1 by RSTAR (= 0.002 +

TAUE) and HBEN, respectively).[2] The desirable characteristic of this estimator is that, by construction, it produces results that seem reasonable. However, this estimator has no known statistical properties and some individuals in each group accept wages below the reservation wage estimates produced. Essentially the same results as described above were found for the relative roles of the arrival rates and acceptance probabilities, i.e., the former continued to dominate in producing the pattern of observed variation in transition rates across groups. As for the levels of the acceptance probability, these were lower for both the gamma and normal distributions. In particular, the levels under the gamma dropped to roughly two-thirds[3]

I also experimented with disaggregation. In particular, I partitioned the sample of white workers by education level.[4] As above, the same results appeared in terms of the relative roles of the arrival rates and acceptance probabilities. The variation in the transition rate across groups continued to appear to be directly related to variation in the arrival rate. Specifically, high education groups have higher arrival rates, which pushes up their transition rates. This finding is consistent with those of Wolpin and Stern for young white males.

---

[2]This is a strategy similar to Narendranathan and Nickell (1985) and Ridder and Gorter (1986). Both of these papers are discussed in detail in Devine and Kiefer (1988).

[3]I describe this as experimentation because convergence was not obtained for all groups.

[4]This disaggregation was done only for the white worker sample because of very small sample sizes in a partition of the sample of nonwhite workers.

Table F.1

IMPLIED DISCOUNT RATES AND NET UNEMPLOYMENT INCOME LEVELS

GROUP:	HBEN[a]	TAUE	RSTAR[b]	IBEN1[c]	IBEN2	DRATE1[d]	DRATE2
White Males:							
Ages 16-19	2.257	0.086	0.088	-1.530	-1.329	-17.839	-18.8
Ages 20-24	3.318	0.059	0.061	-4.197	-4.091	-14.033	-14.2
Ages 25-44	3.498	0.051	0.053	-7.061	-6.223	-23.018	-26.7
Ages 45-64	3.728	0.058	0.060	-4.091	-3.371	-12.292	-14.1
Nonwhite Males:							
Ages 16-19	.	0.083	0.085	1.145	1.382	.	.
Ages 20-24	6.927	0.061	0.063	-2.295	-1.974	-3.689	-3.6
Ages 25-44	2.671	0.046	0.048	-4.462	-2.683	-18.417	-38.9
Ages 45-64	2.776	0.043	0.045	-0.575	-0.014	-25.572	-104.2
White Females:							
Ages 16-19	4.172	0.079	0.081	-1.321	-1.119	-5.984	-6.0
Ages 20-24	2.297	0.068	0.070	-2.048	-2.029	-17.680	-17.8
Ages 25-44	3.232	0.059	0.061	-3.060	-2.950	-11.151	-11.3
Ages 45-64	3.487	0.051	0.053	-1.601	-0.920	-7.254	-8.2
Nonwhite Females:							
Ages 16-19	.[f]	0.103	0.105	1.018	1.421	.	.
Ages 20-24	3.124	0.073	0.075	2.243	2.430	-45.645	-1064.7
Ages 25-44	3.189	0.056	0.058	-1.574	-1.490	-10.242	-10.4
Ages 45-64	2.358	0.047	0.049	-0.222	0.537	-7.422	-12.5

[a,f] HBEN is the group mean for the weekly UI benefit divided by hours on the accepted job. A missing value is indicated by . when no UI benefits are observed for the group sample.

[b] RSTAR is the estimate for $r_c^* = r_c + a_c$. It is the sum of TAUE, the maximum likelihood estimate for a_c, and an estimate of 0.002 for the weekly rate of time preference r_c.

[c] IBEN1 is the hourly net unemployment income level implied by the optimality condition using WR1 for w_c^r and RSTAR for r_c^*. IBEN2 is the same under WR2.

[d] DRATE1 is the weekly discount rate implied by the optimality condition with w_c^r for WR1 and HBEN for b_c. DRATE2 is the same under WR2.

Table F.2 Unemployment Insurance Data

Variable	N	Mean	Standard Deviation
Average Weekly State UI Benefit	966	117.59	97.48
Average Hourly State UI Benefit	966	3.33	3.60