Labour Market Structure and Inequality: A Comparison of Italy and the U.S.

CHRISTOPHER J. FLINN
New York University

First version received April 1997; final version accepted September 2001 (Eds.)

Markets with rigid labour regulations and centralized wage setting are often thought to be inefficient but egalitarian. Using a model of off- and on-the-job search and event-history, individual-level data for Italy and the U.S., we show that while the cross-sectional wage distributions of young Italian males are much more compressed than are the comparable distributions for young white U.S. males, the estimated search model implies that the distribution of lifetime welfare is no more disperse in the U.S. than it is in Italy. Our model implies that the high frequency of movements between labour market states leads to both a relatively equitable distribution of “long run” welfare in the U.S. and a high level of cross-sectional inequality.

1. INTRODUCTION

Institutional environments are often compared in terms of the distributions of individual and family earnings and income which are observed within them. For example, there exists a large literature which contains descriptions of income or earnings inequality across countries, some recent examples of which include Levy and Murame (1992), Gottschalk (1993), Gottschalk and Smeeding (1997) and Atkinson et al. (1992). Typically, normalized measures of dispersion such as the coefficient of variation or Gini index are computed using cross-sectional earnings or income distributions for a number of countries. In this research we make the point that the ranking of institutional environments using inequality measures computed from cross-sectional earnings distributions may be very different from the ranking obtained when lifetime measures of inequality are used.

In comparative studies of cross-sectional earnings or wage inequality, such as those cited above, it is often noted that wage inequality is substantially greater in the U.S. than in most European countries. The histograms displayed in Figures 1(a) and (b) are consistent with this observation. The distribution displayed in Figure 1(a) is of hourly pay rates (measured in 1000s of lira) of young men employed in January 1988 who were living in the Italian region of Lombardia (in which the commercial center Milan is located). Figure 1(b) presents the distribution of hourly wages (measured in dollars) of young white men employed in the U.S. at the same moment in time who were members of the nationally-representative portion of the National Longitudinal Survey of Youth (NLSY). The U.S. wage distribution appears to be more right-skewed than the comparable distribution for Lombardia.1 Using one commonly-utilized inequality measure, the squared coefficient of variation ($C^2$), inequality in the U.S. wage distribution is several times larger than in the corresponding distribution for Lombardia.

1. The right-skew observed in the U.S. data would be even more pronounced if two extremely high wage observations were not first deleted from the U.S. sub-sample we extracted. Based on the demographic characteristics of the individuals involved, we decided that the wage draws were “unbelievably” high.
In addition to their cross-sectional wage distributions, the two samples differ markedly in terms of the rates of transitions across labour market states. The data available to us consist of 17 months of continuous observations on the labour market experiences of sample members in both of the countries, where the observation period runs from January 1988 through May 1989. For the Lombardian sample, 80.8% held the same job over the entire 17 month period, while only 55.6% of the U.S. sample were continuously employed at the same firm throughout. The percentage of sample members who held two or more jobs over the 17 months was 5.2 for Lombardia in contrast with 36.2 for the U.S. Over this observation period 15.4% of the Lombardian sample experienced some nonemployment while fully 28.4% of the U.S. sample was without a job at some point. On the other hand, long-term unemployment is more common in the Italian sample, with 2.9% of the sample reporting unemployment for the entire 17-month period in comparison with less than 1% for the U.S. The overall impression one gets from these figures is that there is much more “churning” in the U.S. labour market than in Lombardia, which is not inconsistent with conventional wisdom.

2. This amount of mobility is not atypical of what is found in different but comparable samples. For example, using the March Current Population Surveys of 1988 and 1989, I was able to match a total of 566 individuals who were white males of the same age as respondents in the NLSY during this period. While it is not possible to determine whether an individual is at the same job in March 1989 as he was in March 1988, the proportion of individuals employed at both points in time whose jobs are at firms in different three-digit industry classifications should provide a lower bound on the annual mobility rate. Of the 453 individuals employed at both survey dates, 66% were working in the same three-digit industry category at both points in time. Since our sample period covers 17 months as opposed to 12, and since moves across firms within the same three-digit category are not recorded, this mobility rate seems consistent with what we observe in the NLSY.
We use the limited wage and turnover information available to us in the two samples to compute measures of lifetime welfare in the two labour markets. To accomplish this, we utilize a search-theoretic framework within which individuals are assumed to act as expected wealth maximizers. Using estimates from this model, we are able to construct lifetime welfare distributions for the two countries. These distributions are plotted in Figures 1(c) and (d). Both distributions are much more symmetric than are the corresponding wage distributions (this is a central limit theorem type of effect), and most strikingly we see that measured inequality in lifetime welfare is now somewhat less for the U.S. than for Italy in terms of the associated squared coefficients of variation.

We will argue, as did Friedman, that the lifetime inequality measure is the most salient indicator of labour market inequality. As he states (1962, p. 171):

A major problem in interpreting evidence on the distribution of income is the need to distinguish two basically different kinds of inequality; temporary, short-run differences in income, and differences in long-run income status. Consider two societies that have the same distribution of annual income. In one there is great mobility and change so that the position of particular families in the income hierarchy varies widely from year to year. In the other, there is great rigidity so that each family stays in the same position year after year. Clearly, in any meaningful sense, the second would be the more unequal society. The one kind of inequality is a sign of dynamic change, social mobility, equality of opportunity; the other, of a status society.

The inequality analysis we conduct, both in terms of short-run and long-run welfare distributions, allows us to decompose inequality within each environment into within- and between-group components, where in this analysis groups are defined in terms of years of schooling completed and/or educational credentials. This allows us to investigate the extent to which differences in inequality between Lombardia and the U.S. reflect differences in inequality within broadly-defined schooling groups and to what degree they are produced by between-group differences in welfare. We may find that in the aggregate, environment A is less “unequal” than environment B, but that within environment A all inequality is between schooling groups [i.e. within schooling groups all individuals have the same level of welfare] while in B there exists no between-group inequality. In this case, while overall inequality may be less in A, inequality in A may be said to be “predetermined” in that where one ends up in the welfare distribution depends solely on one’s schooling level. Without an explicit model of the manner in which schooling choices are made it is not possible to say whether or not “predetermination” is good or bad from a fairness perspective. Nevertheless, the proportion of inequality attributable to between-group differences in welfare is informative at a descriptive level on the structure of rewards within an environment.

We have chosen Italy and the U.S. to make our substantive points because the labour market institutions in these two countries are well-known to be quite dissimilar. Our analysis allows us to identify the effects of characteristics of the labour market environment, such as rates of receiving offers while unemployed or employed, the dismissal rates, and the wage offer distribution, on the distribution of life-cycle welfare. However, since we do not estimate a general equilibrium model of the labour markets in Italy and the U.S. we will not be able to claim that differences in parameter estimates we obtain and outcome distributions derived from them are strictly attributable to the institutional features of the two countries.3

There exist a number of studies which compare the labour markets of Italy and the U.S., both in terms of institutional differences and empirically (see, for example, Grubb and Wells

3. For an attempt to conduct a comparative analysis within a partial equilibrium model see Cohen (1999), who compares labour market outcomes in France and the U.S. using an insider-outsider framework.
(1993), Del Boca (1988), Bertola and Ichino (1995), Bertola and Rogerson (1997), Katz et al. (1995), Demekas (1994)). In virtually all of the empirical studies, differences in the industrial relations systems or output markets are invoked to explain patterns of intertemporal relationships between comparable (typically aggregated) labour market measures in the two countries. A popular focus of this type of analysis is the role of institutional restrictions on layoffs and dismissals in accounting for the lower level of cyclical variability in employment levels observed in Italy. Demekas (1994) presents a very useful survey of the differences in the industrial relations systems of Italy and the U.S. at the time in which the data analysed here was collected.

There are other approaches to the analysis of long-run welfare inequality which can be profitably taken. One alternative is strictly empirical, and consists of analysing the distribution of time-aggregated earnings and income data (see, for example, Aaberge et al. (1996), Björklund (1993), Blomquist (1981)). This approach has as its main advantage the fact that no modelling framework, with its attendant assumptions, need be adopted to compare inequality in different institutional environments. This is a very great advantage, especially when one attempts to compare compensation distributions across institutional environments in which individual behaviour takes place in qualitatively different contexts. The main disadvantages of such an approach are that the requisite data for performing it are only available for a few populations, and the absence of an explicit modelling framework makes it difficult to generalize results to other cohorts and to determine which features of the institutional environment are important in determining the long-run welfare distribution.

A variant of this approach utilizes subjective information gathered from respondents regarding their perception of the likelihood of future labour market events. In this manner, the need for extensive amounts of retrospective information on the labour market experiences of mature workers is alleviated. The fact that respondents are asked to report expectations concerning future events means that it is not possible to validate the information provided, although the use of such information obviates the need for assumptions regarding individual information sets and decision-making processes. The principle disadvantage, in addition to the usual issues which arise in analysing subjective information, is the inability to perform policy experiments. Good examples of this type of analysis are Guiso et al. (1998) and Dominitz and Manski (1997). To my knowledge, these types of data have not been used in an explicitly comparative analysis.

A third approach is to change the focus of the analysis so as to explicitly analyse income or earnings dynamics outside of a specific modelling framework. The objective of this type of study is, treating compensation over the life-cycle as an exogenous stochastic process \( F \), to find summary measures characterizing the degree of mobility, \( \sigma(F) \). Several mobility measures have been proposed by economists and sociologists, and there exist a number of empirical studies in which the mobility process is the focus of interest.\(^4\) The advantage of this approach is that it is explicitly dynamic and can provide useful summary information regarding how much mobility exists in an environment—the summary measures produced are explicitly designed for comparative analysis in that they are unitless. The principle disadvantage is in not being able to explicitly value (in welfare terms) the mobility process associated with a given environment. This is only to say that static distributional and mobility analysis are only pieces of the puzzle—neither is sufficient in and of itself to determine the distribution of life-cycle welfare.

Related to the first and third strategies is one in which earnings, income, and/or consumption processes are decomposed into permanent and transitory components which are posited to have

---

\(^4\) In terms of the derivation of mobility indices, some examples of influential work on the subject are Prais (1955), Shorrocks (1978a,b) and Geweke et al. (1986). For a recent example of empirical work focused on earnings mobility see Atkinson et al. (1992).
different implications for long-term and short-term welfare distributions. In some cases the analysis is primarily descriptive (e.g. Gottschalk and Moffitt, 1994), while in others econometric identification is obtained by linking the observed processes through the specification of a utility function and optimization problem (Blundell and Preston, 1998). This last paper is perhaps closest in spirit to the analysis we have performed here in its emphasis on the differences between short-run and long-run welfare distributions. The agents in the Blundell and Preston analysis are assumed to be risk-averse, which is perhaps more realistic than the assumption of risk-neutrality that we impose. The practical advantages of the risk neutrality assumption are the simple form of the implied decision rules used by agents and the ability to estimate the model with what is essentially cross-sectional information.

Ours is certainly not the first analysis to utilize a search-theoretic framework to analyse differences in outcomes between observationally distinguishable groups. For example, Wolpin (1992) uses a model of off-the-job search to analyse racial disparities in unemployment and wages. He utilizes a discrete time framework, and only considers the school-leaving, first-job taking process. Eckstein and Wolpin (1995) look at these same phenomena in the context of a matching–bargaining framework to isolate the rate of return on schooling investment (as reflected by shifts in the wage offer distribution) from other impacts on parameters characterizing the search technology. Using aggregate labour market data from different countries, our approach is differentiated from these and other empirical studies of this type primarily by its emphasis on inequality of wage and welfare distributions. Since utility levels are not directly comparable across institutional environments, our comparative analysis must be based on unitless measures such as inequality indices.

The plan of the paper is as follows. In Section 2 we describe the data sources and provide further descriptive evidence on labour market dynamics at the individual level in the two markets. We also discuss the manner in which inequality is measured and the "analysis of variance" types of decomposition into between- and within-group components we are able to perform. Section 3 provides the theoretical structure for our analysis; in it we define the labour market environment and the decision rules used by participants. Section 4 contains a discussion of econometric issues and the maximum likelihood estimator we utilize. We present estimation and simulation results in Section 5, and offer a brief conclusion in Section 6.

2. DATA AND MEASUREMENT

In this section we describe the data to be used in all of the empirical analysis which follows, and then present some descriptive statistics which further serve to illustrate the marked differences in transition rates between labour market states in these two environments. The section concludes with a development of the measures of inequality we employ.

2.1. Data description

We begin by describing the event-history data for the U.S. labour market, which is taken from the random sample respondents of the NLSY. This data set is very well-known to most labour market researchers, so only a brief discussion of the information we have extracted and the sample-selection criteria will be presented.

As we describe below, the Italian data utilized come from a random sample of households in the Lombardia region of Italy. For such a large industrial and commercial area, the labour force is relatively homogeneous in the demographic sense. For purposes of comparison, we have restricted the U.S. sample to include only whites in an attempt to limit the amount of sample heterogeneity in background characteristics. Furthermore, since the observation period for the
Lombardian data only covers the period January 1988 through May 1989, we have constructed a continuous labour market history from the NLSY data that covers only this period as well. Since the NLSY sample members were 14–21 years of age in the initial survey year of 1979, they are (approximately) between the ages of 24–31 during the observation period. For purposes of comparability, we have only utilized individuals from the Italian sample who were between 24–31 during the observation period. We have further restricted the samples to include only males, not only for the usual reason of not wishing to confront the problem of modelling the participation decisions of married women, but also because the patterns of female labour force participation are so different in the two environments.5

For this sample of white males we have constructed event histories which include information on the duration of each job held, where a “job” is defined as an employment spell spent with a particular employer, and the hourly wage rate received on the job; we also have information on the length of time spent in each spell of nonemployment. We do not attempt to distinguish between periods spent actively searching and those spent out of the labour force—all are aggregated into the category “nonemployment” (though to avoid redundancy we shall sometimes call this state “unemployment”). We have excluded from our final sample all individuals who reported never actively searching for employment or who did not hold a job over the observation period. Thus everyone in the final sample is a labour market participant in the formal sense of the term at some point in the 17-month period. We have also excluded individuals who reported being enrolled in school or in the military during the observation period, as well as all those remaining individuals who were not in the random sample component of the NLSY. Due to the difficulties of defining an appropriate compensation measure, we have also excluded sample members who reported being self-employed at some time over the period.6

The Italian Statistical Institute (ISTAT) survey was conducted in Lombardia during the summer of 1989, and gathered information from a sample of households (as opposed to the NLSY, in which the sampling unit is the individual). Respondents were asked to describe their labour market experiences rather completely over the 17-month period that spanned January 1988 through May 1989. In particular, respondents reported their labour market status on a month-by-month basis over this period. They were then asked to provide detailed information on up to two jobs held as a dependent worker and one job as an independent worker during this period. All duration information was given solely on a monthly basis.

In Italy respondents in social surveys are infrequently asked for information relating to their earnings or income from other sources. Given that wage and/or earnings information is an essential requirement for estimating a search model, this is the only Italian survey information in existence which could be used for this purpose.7 Respondents were asked, as in the NLSY, to

5. It could be argued that patterns of household inequality in labour market outcomes are ultimately of more interest than are patterns of individual-level inequality. However, to perform a comparative analysis at this level requires a model of household formation and household decision-making that is well beyond the scope of this paper.

6. This was the last sample selection criterion imposed in defining the final sample for both markets. In the case of Lombardia, this criteria resulted in the exclusion of 15.4% of the remaining sample. Excluding respondents who were self-employed at some point in the 17-month period in the NLSY resulted in a loss of 9.8% of the cases. This appreciable difference in self-employment rates in the two countries led us to develop an extended version of the single-sector search model used in the present analysis that includes a self-employment option. In that model, workers can move between dependent worker and self-employment status as options in both sectors become available at times generated by independent Poisson processes. We estimated that model using data from the Lombardia sample, and found that the impact on estimated life cycle welfare distributions was not substantial. Details are available from the author upon request.

7. There are some administrative records from the Italian social security administration (INPS) which could possibly be used, but it is difficult to ascribe earnings to a particular job using these data. In addition, only jobs covered by social security provisions are included; for example, all public sector jobs are excluded. The Bank of Italy panel data set also has problems of time aggregation that make its use extremely problematic for our purposes.
FLINN LABOUR MARKET STRUCTURE

report a typical level of earnings for each job occurring during the observation period. Thus the duration and remuneration information collected is quite similar in the two surveys.

We conclude this discussion with four remarks on the use of these samples for comparative analysis. First, labour market dynamics at the individual level, especially for young labour market participants, are likely to be quite sensitive to aggregate labour market conditions. In order to perform a comparative analysis which heavily relies on stationarity assumptions in order to extrapolate labour market dynamics from a brief historical period to the entire life-cycle, it is at a minimum necessary to ensure that the two economies were not at very different stages in the business cycle during our observation period. Fortunately for us, in 1988 and 1989 the U.S. and Lombardia seem to have been at similar stages in the business cycle. We arrived at this judgement by computing standardized values of annual gross domestic product in constant local prices for Lombardia and the U.S. over the period 1980–1993. The values of this standardized measure in 1988 and 1989 were 0.602 and 0.982 in Lombardia in comparison with 0.568 and 0.828 in the U.S. When we computed these measures after first detrending each series, we arrived at values of 0.918 and 1.452 for Lombardia and 1.221 and 1.359 for the U.S. On the basis of either set of values both economies seem to have been in relatively “good” states.

Our second remark concerns the treatment of within-country heterogeneity in the analysis. We made a decision early on not to introduce an extensive set of conditioning variables into either the modelling framework or the empirical analysis. When allowing individuals in each country to be observationally differentiated from one another, groups must be defined within each country, and defining “comparable” partitions across countries is inherently difficult. The definition of within-country partitions can affect our inferences regarding within and between group inequality for each labour market and the comparison of these breakdowns across labour markets. In the end, we have decided to distinguish a small number of education classes within each country, and to perform an inequality decomposition using these schooling groups. In addition to allowing for observable heterogeneity, it is in principle possible to allow for unobservable heterogeneity as well. We attempted to do this but, given the short observation period available to us and relatively small sample sizes, it proved impossible to estimate the requisite behavioural models. We provide further discussion of this issue in Section 4.3.

A third issue relates to the treatment of wage variability within job spells. While in actuality wages are variable over job spells in both labour markets, we have little hope of being able to precisely estimate within-spell wage processes in either sample. For Lombardia we have no information on wage changes within job spells since only one compensation measure is ascertained per spell. Because the NLSY is a panel, for jobs held across sampling times (separated by approximately one year) the contemporaneous wage rate (i.e. at the time of the survey) is recorded. While such information could in principle be used to estimate an “on-the-job” wage process, the job spells for which this information is available clearly do not constitute a random sample from the population of all job spells. Estimation of population parameters in such a case would be extremely problematic. As a result of these data limitations we are not able to consider the contribution of within-job spell wage variability to the total variability observed in the market at a point in time or to differences in life-cycle welfare in either country.

Finally, we are obliged to point out a problem of potential noncomparability with regards to the wage measures used in this study. The ISTAT survey in Lombardia explicitly asks respondents to report their net (i.e. after-tax) monthly earnings. Hourly wages are imputed by dividing this measure by the reported usual hours worked at the employer. On the other hand, in the NLSY individuals are asked to report how much they earned when they first began working at the employer. The respondent is not told to report gross (i.e. pre-tax) or net earnings, so that we can only speculate as to how respondents interpreted the question. Anecdotal evidence suggests that respondents in the U.S. are more apt to report gross earnings when asked about compensation,
while Italians are more likely to report net earnings. If members of the U.S. sample report gross earnings, how is this likely to affect the comparative analysis? The answer depends in part on the nature of the tax system. If it is proportional or progressive, the gross wage distribution will be more disperse than the net wage distribution. In this case, cross-sectional differences in the wage inequality in Italy and the U.S. may be over-stated. Since the principle message of the paper is that lifetime welfare differences in the U.S. labour market may be no larger than those which exist in Lombardia, we would argue that having comparable measures (say net wages) would probably serve to further strengthen these findings. Unfortunately, we cannot provide any further evidence on this issue with the data at our disposal.

2.2. Variable definitions and descriptive statistics

The wage variables used in the analysis were described above. Since the samples are relatively homogeneous with respect to all characteristics other than schooling attainment, we have only to describe our definitions of schooling groups.

We defined three schooling categories for Lombardia and four for the U.S. Within each country the categories can be thought of as running from “low” schooling attainment to “high” schooling attainment, though the categories themselves cannot legitimately be compared across countries. The classification used for the U.S. is relatively standard: the low schooling group consists of individuals who did not complete high school, the second is comprised of high school graduates with no further education, the third consists of individuals with a high school diploma who attended college but did not earn a (4-year) college degree, and the last consists of individuals who have earned at least a (4-year) college degree. For Italy, the lowest group is comprised of individuals who have completed middle school or a high school programme of less than 4 years, the second group consists of individuals who have completed a “full” high school programme, and the last group consists of those who hold a university degree.8 Note that the second group, those who have completed a high school programme of 4 or 5 years will also include many individuals who have completed some college courses. In this sense, the Italian category S2 is most similar to the U.S. categories S2 and S3:

<table>
<thead>
<tr>
<th>Category</th>
<th>Lombardia</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>Elementary School Certificate</td>
<td>Less than H.S. Graduate</td>
</tr>
<tr>
<td>S2</td>
<td>Middle School Certificate</td>
<td>H.S. Graduate</td>
</tr>
<tr>
<td>S3</td>
<td>High School Diploma</td>
<td>Some College</td>
</tr>
<tr>
<td>S4</td>
<td>Laurea</td>
<td>College Grad and Above</td>
</tr>
</tbody>
</table>

Table 1 contains some descriptive statistics for our samples of U.S. and Lombardian labour market participants. First note that there are approximately four times as many individuals in the U.S. sample. In terms of numbers of spells, the discrepancy is even larger since transitions between labour market states are much more frequent in NLSY than in the ISTAT sample.

The age distributions are extremely similar, and this results in the virtual equality of the mean and standard deviation of age in the two samples. Although schooling groups cannot be compared across the two environments, we note that the distribution is much more concentrated

8. For those readers familiar with the Italian educational system, category S1 is comprised of individuals with a licenza di scuola elementare or licenza di scuola media inferiore, S2 includes those with a diploma or licenza di scuola media superiore, and S3 includes those with a laurea or diploma universitario. The second category includes individuals who have completed a programme of secondary studies that does not enable them to enroll in a university.
FLINN  LABOUR MARKET STRUCTURE  619

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Lombardia</th>
<th>U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size (individuals)</td>
<td>422</td>
<td>1518</td>
</tr>
<tr>
<td>Age in 1989</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>28.406</td>
<td>28.246</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.305</td>
<td>2.209</td>
</tr>
<tr>
<td>Proportion of sample members with:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Schooling</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>0.588</td>
<td>0.174</td>
</tr>
<tr>
<td>S2</td>
<td>0.336</td>
<td>0.327</td>
</tr>
<tr>
<td>S3</td>
<td>0.076</td>
<td>0.212</td>
</tr>
<tr>
<td>S4</td>
<td>NA</td>
<td>0.287</td>
</tr>
<tr>
<td>Some nonemployment</td>
<td>0.154</td>
<td>0.284</td>
</tr>
<tr>
<td>All nonemployment</td>
<td>0.028</td>
<td>0.009</td>
</tr>
<tr>
<td>Same job all 17 months</td>
<td>0.808</td>
<td>0.556</td>
</tr>
<tr>
<td>Two or more jobs</td>
<td>0.052</td>
<td>0.362</td>
</tr>
<tr>
<td>Wages on all jobs</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>7.835</td>
<td>8.695</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>2.543</td>
<td>7.083</td>
</tr>
<tr>
<td>Two job spells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>12</td>
<td>402</td>
</tr>
<tr>
<td>Proportion with wage gain</td>
<td>0.917</td>
<td>0.711</td>
</tr>
</tbody>
</table>

In Lombardia than in the U.S. In part, this is an artefact of the coding employed by ISTAT, but it is of substantive interest to note that relatively few Italians have completed a university degree. In fact, in this sample only 7% held a college degree (a laurea), compared to the almost 29% of the U.S. sample that hold at least a Bachelor's degree. This difference is partially due to the fact that the laurea requires more years to complete and is, very generally speaking, a more specialized degree. The category "some college" in the U.S. contains 21.2% of the sample and those who have finished high school and did not attend college comprise 32.7% of the sample. As argued above, the category S2 for Italy is roughly analogous to the aggregate of S2 and S3 for the U.S., and contains 33.6% of the sample. The lowest schooling category for the U.S. contains only 17.4% of the sample, while for Lombardia the majority (58.8%) of sample members have not completed a full high school programme.

We have already commented upon the large differences in descriptive statistics regarding labour market transitions in the two markets in Section 1. Using the units of dollars per hour for the U.S. and thousands of lire per hour for Lombardia, the mean wage rates are not too dissimilar, however the standard deviation of wages in the U.S. is almost three times as large as is the one for Lombardia. This similarity in means and appreciable difference in standard deviations produced the large discrepancy in C^2 values reported in Figures 1(a) and (b).

The model of on-the-job search estimated in what follows implies that direct job-to-job transitions will always be associated with a wage increase. In the Italian sample, no individual reported having more than two jobs over the 17-month period. Of the 16 who reported having two jobs, 12 reported that the jobs were consecutive. Of these 12, the wage on the new job was greater than the wage on the first in 11 cases.

For the U.S. sample, 535 individuals reported having two or more jobs. Within this set of individuals, 389 individuals had two or more consecutive jobs, and some had multiple consecutive job spells within the 17-month period. The total number of employment spells
containing two or more consecutive jobs is 402. Of these 402 spells, the second job in the spell paid a wage greater than that paid by the first job in 286 cases, yielding a “success rate” of 0.711. Thus in both samples, the tendency to have higher wages in the second job is clearly present, though, especially in the U.S. sample, it is far from always the case. When we discuss estimation issues below we will consider ways in which the search model can be modified so as to account for these discrepancies.

2.3. Measuring inequality

In analysing inequality within and across the two countries we will utilize three standard measures. The first is the squared coefficient of variation, which we have denoted by $C^2(y)$, where the argument $y$ is an $n \times 1$ vector of outcome measures $(y_1 \ y_2 \ \ldots \ \ y_n)'$ for the characteristic of interest. The measure is defined as

$$C^2(y) = \frac{n^{-1} \sum_{i=1}^{n} (y_i - \bar{y})^2}{\bar{y}^2},$$

where $\bar{y} \equiv n^{-1} \sum_{i=1}^{n} y_i$. The second measure utilized is Theil’s entropy measure, $T(y)$, which is defined by

$$T(y) = \sum_{i=1}^{n} \frac{y_i}{Y} \ln \left( \frac{y_i/Y}{1/n} \right),$$

where $Y \equiv \sum_{i=1}^{n} y_i$. The final measure which we consider is Theil’s second entropy-based index, $L(y)$, which is given by

$$L(y) = \sum_{i=1}^{n} \frac{1}{n} \ln \left( \frac{1/n}{y_i/Y} \right).$$

Clearly, Theil’s two measures are closely related. In particular, the “roles” of $n^{-1}$ and $y_i/Y$ (which is individual i’s share of the total pie, $Y$) are reversed.

All three measures satisfy the highly desirable properties of (i) mean and scale independence (that is, $M(y) = M(a + by)$); (ii) population-size independence ($M(y) = M((y_1' \ldots \ y_k')'$), $k = 1, 2, \ldots$); and (iii) satisfaction of the Pigou–Dalton condition, which states that whenever a “better-off” individual $i$ makes a transfer to a worse-off individual $j$ [i.e. $y_i > y_j$] of $t < y_i - y_j$, the resulting distribution should be less unequal than the original. Especially important for our purposes is the ability to meaningfully decompose each index into inequality between and within identifiable subpopulations. In our application we will be interested in determining the amount of inequality within and between schooling groups in each country. Consider the case in which there are $S$ subpopulations in a particular country differentiated by schooling level, and let the vector of values of $y$ in subgroup $s$ be denoted $y(s)$. Then if $M$ is an additively decomposable index it can be written as

$$M((y(1)' \ldots y(S)')') = \sum_{s=1}^{S} \theta_s M(y(s)) + M((\bar{y}(1)' \ldots \bar{y}(S)')'),$$

where $\bar{y}(s)$ denotes the vector created when the value of $y_i$ for each member of group $s$ is replaced with the average value of $\bar{y}_s$ for the group. The weights $\theta_s$ differ across the three

9. This section largely follows the excellent development of this material in Anand (1983).
inequality measures; in particular we have

<table>
<thead>
<tr>
<th>Index</th>
<th>$\theta_s^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C^2$</td>
<td>$\frac{n_s}{n} \left( \frac{\overline{y}}{\bar{y}} \right)^2$</td>
</tr>
<tr>
<td>$T$</td>
<td>$\frac{n_s}{n} \frac{\overline{y}}{\bar{y}}$</td>
</tr>
<tr>
<td>$L$</td>
<td>$\frac{n_s}{n}$</td>
</tr>
</tbody>
</table>

Note that the “within” weights associated with the measures $C^2$ and $T$ both depend on the distribution of $y$ as well as the subgroup sizes, $n_s$. Decompositions based on the index $L$ are particularly advantageous since they are independent of the subgroup mean values $\overline{y}_s$.

Decompositions made using these measures can be employed to determine the amount of inequality due to differences within and between subgroups in a manner similar to the decomposition of differences in levels between and within groups using the analysis of variance. In particular, besides comparing inequality levels within subgroups, to a limited extent we can measure the proportion of total inequality due to differences across subgroups. The proportion of total inequality (according to any one of the three measures) which is accounted for by between schooling group differences is given by

$$IW = \frac{M((\overline{y}(1) \ldots \overline{y}(S))')}{M((y(1) \ldots y(S))')}.$$ 

In terms of cross-sectional wage rates, we have already seen in Figure 1 that the U.S. sample displays considerably more inequality than does the Italian (Lombardian) sample; this is the case no matter which of the three inequality measures is used. In all cases, the U.S. measure is approximately three times larger than the corresponding Italian one.

Table 2 contains decompositions of the cross-sectional wage distributions using the three inequality measures. Within the U.S. sample, some interesting relationships between schooling levels and within-group inequality emerge. Inequality seems to be most pronounced within the group of individuals who have attended but not completed college ($S_3$). On the other hand, the most homogeneous group appears to be those who have (at the minimum) completed college ($S_4$). It may be the case that the college-educated group is the most homogeneous because its members are the most recent labour market entrants and individual differences in earnings have not yet been manifested. The college-attending group may exhibit substantial inequality because its members can be expected to be quite heterogeneous, consisting of individuals who have completed 2 year vocational degree programmes, individuals who have dropped out of 4 year programmes to take a good job, and others who have left due to poor performance in school. The two lower schooling groups have similar levels of inequality in wages, though the group which completed high school ($S_2$) has slightly more unequal wages than the drop outs. The proportion of total inequality attributable to between schooling group differences is only about 5 or 6% (depending on the inequality measure used). Thus the vast majority of inequality within the U.S. sample is not attributable to schooling differences.

For the Lombardian sample the situation is a bit different. The small group of individuals ($S_3$) with university degrees exhibits a relatively high level of wage inequality, although it is small in comparison with the amount of inequality in the highest schooling group in the U.S. sample, or any of the other schooling groups for that matter. The lowest schooling group is also relatively heterogeneous, though this is to be expected given the large amount of aggregation performed in defining it. The middle schooling group has a far lower degree of inequality than either of the other two schooling groups, which is a bit surprising given that it consists also of individuals with some college experience and therefore could be expected to be heterogeneous.
As we have seen, the total amount of inequality in current period wages is far less in Lombardia than in the U.S. However, the proportion of inequality attributable to differences in educational levels ($IW$) is much higher in Lombardia than in the U.S. $IW$ ranges between 10 and 12% for the Italian sample, whereas it is only 4 to 6% for the U.S. In terms of wages, schooling differences account for a much larger proportion of the total inequality in Lombardia, although in absolute terms within-schooling group inequality predominates in both environments.

3. THE BEHAVIOURAL MODEL

In this section we set out the behavioural model used in performing the empirical work below, which is one of off- and on-the-job search in a stationary labour market environment. We assume that agents are infinitely lived (or are finitely lived and face a constant risk of death) and that at each moment in time they occupy one of the following labour market states. They may be nonemployed and searching for work, a state denoted by $n$, or they may be employed in a job which pays a wage $w$—this state will be denoted $e(w)$. Job offers arrive to nonemployed individuals according to a Poisson process with parameter $\lambda_n$, and to employed individuals according to a Poisson process with parameter $\lambda_e$. For both employed and nonemployed searchers the value of the offer made is independent of the arrival process. Employment spells terminate either because (1) an employed searcher locates a better job or (2) an exogenous separation occurs. Exogenous separations arrive according to a Poisson process with parameter $\eta$; note that the rate of “involuntary” separations is assumed to be independent of the wage paid on the job. The instantaneous rate of discount is $\rho$. Nonemployed searchers receive an instantaneous net benefit of $b$ which is of unrestricted sign. The wage offer distribution is given by $F$ (and the associated survivor function is denoted $\bar{F} = 1 - F$); the successive offers received by an individual are taken to be independently and identically distributed (i.i.d.) draws from $F$. 
Consider an infinitesimally small period of time $\Delta t$. An agent who is currently employed at the instantaneous wage rate $w$ may find himself in one of three states at the “end” of interval $\Delta t$. First, nothing may happen during the interval so that the individual remains employed at wage $w$. Second, the agent may be exogenously separated from his or her job and thus enter state $n$ with associated value $V_n$. Finally, the agent may receive a new job offer with associated wage rate $w'$. Since the wage is the only characteristic of the job valued by the individual and there are no fixed costs associated with job-changing, the individual will change jobs if and only if $w' > w$, in which case the end of “period” value will be $V_e(w')$. If $w' \leq w$ the agent will reject the offer and stay with his current employer; in this case the end of period value of employment remains $V_e(w)$. Thus the value of employment over the interval $\Delta t$ is given by

$$V_e(w) = (1 + \rho \Delta t)^{-1} \{ w \Delta t + \eta \Delta t V_n + (1 - \eta \Delta t - \lambda_e \Delta t) V_e(w) + \lambda_e \Delta t \int \max[V_e(w), V_e(x)] dF(x) + o(\Delta t) \},$$

where the term $(1 + \rho \Delta t)^{-1}$ is the “infinitesimal” discount factor associated with interval $\Delta t$, $\eta \Delta t$ is the approximate probability of being terminated from one’s current job at the end of the interval $\Delta t$, $(1 - \eta \Delta t - \lambda_e \Delta t)$ is the approximate probability of not being terminated or receiving a new job offer, $\lambda_e \Delta t$ is the approximate probability of receiving a new job offer, $V_n$ is the value of the nonemployment state, and $o(\Delta t)$ is a term which has the property that $\lim_{\Delta t \to 0} o(\Delta t)/\Delta t = 0$. After collecting terms and taking the limit of (1) as $\Delta t \to 0$, we have

$$V_e(w) = \frac{w + \eta V_n + \lambda_e \int V_e(x) dF(x)}{\rho + \eta + \lambda_e F(w)}.$$  

The value of nonemployment is similarly derived. Over a small interval $\Delta t$ the individual can either receive one job offer, with approximate probability $\lambda_n \Delta t$, or none. Then the value of nonemployment over the interval $\Delta t$ is

$$V_n = (1 + \rho \Delta t)^{-1} \{ b \Delta t + \lambda_n \Delta t \int \max[V_n, V_e(x)] dF(x) + (1 - \lambda_n \Delta t) V_n + o(\Delta t) \}. $$

After rearranging terms and taking limits, we have

$$V_n = \frac{b + \lambda_n \int \max[V_n, V_e(x)] dF(x)}{\rho + \lambda_n F(w^*)}.$$  

Given that an offer $w$ is received by a nonemployed individual it will be accepted if and only if $V_e(w) \geq V_n$. Since the value of search is a constant and given the easily demonstrated result that $V_e$ is monotone increasing in its argument, it follows that there exists a unique reservation wage $w^*$ such that $V_e(w^*) = V_n$. We can rewrite (4) as

$$V_n = \frac{b + \lambda_n \int w^* V_e(x) dF(x)}{\rho + \lambda_n F(w^*)}.$$  

The reservation wage $w^*$ is solely a parameter that characterizes the nonemployed searchers decision rule; it is a function of all the “primitive” parameters of the model: $\lambda_e$, $\lambda_n$, $\rho$, $\eta$, $b$ and $F$.

While the state valuation functions $V_e(w)$ and $V_n$ are nonanalytic functions of all the primitive parameters of the model, the decision rules themselves are very simply characterized. If the agent is nonemployed and receives an offer $w$:

$$d_n^*(w) = \begin{cases} \text{accept offer } w & \Leftrightarrow w \geq w^* \\ \text{continue nonemployment} & \Leftrightarrow w < w^* \end{cases}.$$  

This content downloaded from 128.122.230.132 on Tue, 03 May 2016 16:56:40 UTC
All use subject to http://about.jstor.org/terms
If the agent is currently employed at some wage $w$ (greater than $w^*$ by definition) and receives an offer $w'$ his rule is

$$d^*_e(w') = \begin{cases} 
\text{accept new job } w' & \text{ if } w' > w \\
\text{continue current job } w & \text{ if } w' \leq w
\end{cases}$$

(7)

These simple rules greatly facilitate estimation of the behavioural model, as we shall see below.

4. ECONOMETRIC ISSUES

We consider the formulation of a maximum likelihood estimator that consistently estimates virtually all of the primitive parameters characterizing the economic environment. These parameters are of interest in their own right as a basis for comparison of individual level labour market dynamics in the two environments, and perhaps are even more useful in serving as a foundation for the generation of the life-cycle welfare simulations that follow. Before turning to the mechanics of defining the estimator we must first consider three important preliminary issues—initial conditions problems in the data, the modification of the model to account for job-to-job transitions associated with wage decreases, and the introduction of individual-level heterogeneity. While these issues may appear strictly technical on the surface, they actually raise a number of interesting substantive concerns and (hopefully) lead to some insights regarding the interpretation of the model and of cross-sectional remuneration distributions.

4.1. Initial conditions problems

The (limited) retrospective nature of the labour market information collected in the Italian data implies that the spell first sampled, that is the one on-going in January 1988, causes a nontrivial initial conditions problem when attempting to estimate a model of on-the-job search. If an individual was not employed at that time, no pre-sample information was collected. If an individual was currently employed at the beginning of the sample period (January 1988), he was asked the date of the beginning of that particular job spell. The individual was not asked how many jobs he had in succession before the one occupied in January 1988.

The nature of the problem can be illustrated by referring to Figure 2, which displays two hypothetical labour market histories, A and B. Individual A exited from nonemployment into employment at time $r_1$ at a job paying a wage of $w_1$. The wage paid on this first job in the employment spell must exceed the reservation wage $w^*$, so that the conditional distribution of the wage on the first job after nonemployment is $F_1(w) = (F(w) - F(w^*))/F(w^*)$, $w > w^*$, where $F$ is the population wage offer distribution. A found a better job (in the sense of one paying a higher wage, [i.e. $w_2 > w_1$]) at time $r_2$. Individual A continued to find better jobs, and was in his fourth consecutive job in the current employment spell when the observation period began. The job in progress at the time of observation period began at time $r_4$ and paid a wage $w_4$, and we know that $w_4 > w_3 > w_2 > w_1$. Unconditional on the wages associated with previous jobs in this employment spell, the wage associated with the sampled job spell will be a random draw from the distribution $F_4(w)$, which is defined as the wage distribution of individuals who have had exactly three consecutive jobs in the current employment spell prior to the current job. Assuming the wage offer distribution $F(w)$ is continuously differentiable so that the density $f(w)$ exists everywhere, the distribution $F_4$ has an associated probability density function which is given by

$$f_4(w) = \int_{w_4}^{w} \int_{w_3}^{w} \int_{w_2}^{w} \frac{f(w)}{F(w_3)} \frac{f(w_3)}{F(w_2)} \frac{f(w_2)}{F(w_1)} \frac{f(w_1)}{F(w^*)} dw_3 dw_2 dw_1, \quad w \geq w^*.$$  

(8)
In Appendix A we derive an expression for the densities $f_j(w)$, $j = 2, 3, \ldots$. These densities are only slightly special cases of standard order statistic density functions.

In contrast, individual B was in his first job spell after nonemployment when the observation period began. This job began at time $\sigma_1$ and paid a wage $w_1'$, which is an i.i.d. draw from the truncated distribution $F_1(w')$ given above.

In the empirical work conducted below we assume that wage offers are i.i.d. draws from the Pareto distribution, the c.d.f. of which is given by

$$F(w) = 1 - \left(\frac{w^*}{w}\right)^\alpha, \quad w \geq w^* > 0, \quad \alpha > 0.$$ 

and the p.d.f. is given by

$$f(w) = \frac{\alpha (w^*)^{\alpha}}{w^{\alpha+1}}.$$ 

Notice that we have assumed that the lower bound of the wage offer distribution is equal to the reservation wage in the population. Such a result is consistent with the implications of the Burdett–Mortensen equilibrium search model (1998), for example.

The form of the job-spell specific wage densities is particularly simple under the Pareto assumption used in the empirical work. In this case the densities are given by the closed-form expression

$$f_j(w) = f(w) \times \frac{\alpha^{j-1} \ln^{j-1}(\frac{w}{w^*})}{(j-1)!}, \quad j = 1, 2, \ldots.$$  \hspace{1cm} (9)

Figure 3 contains plots of the wage p.d.f.s associated with the first four job spells in any employment spell using parameter estimates for Lombardia and the U.S. Note the large differences in the densities across job spells, perhaps most notably with respect to the degree of symmetry exhibited. Given that these distributions are essentially those of order statistics, they can be ranked in terms of standard stochastic dominance criteria. The figure also serves to illustrate this point.

The transitions between job spells and the unemployment state leads to a steady-state distribution in terms of state occupancy—the probabilities of finding a randomly sampled individual in nonemployment, in the first job of an employment spell, in the second job of an employment spell, etc. Let the conditional probability that an employed individual is in his $i$th
job in an employment spell be denoted by \( q_i \). Then the steady state p.d.f. of wages is of the form

\[
f^S(w) = \sum_{i=1}^{\infty} q_i f_i(w), \quad w \geq w^*,
\]

\[
\sum_{i=1}^{\infty} q_i = 1, \quad q_i \geq 0, \quad i = 1, 2, \ldots
\] (10)

The probabilities \( q \) are derived in Appendix A. It should be clear that they depend on all the structural parameters of the model since the transition rates between states are functions of the parameters characterizing exogenous labour market processes (such as the rates of receiving offers and dismissal) as well as the strategies of the labour market participants.

The steady state wage distribution has two potential uses with regards to this study. First, it is of substantive interest in that it provides a link between a stationary model of on-the-job search and the cross-sectional remuneration distribution that is a focus of much work on income inequality. In particular, it provides us with a good example of the interpretative problem to which Friedman was referring in the citation provided in the introduction. For consider increasing the rate of receiving offers while employed, \( \lambda_e \), in an on-the-job search environment. This will increase the value of all states, \( n \) and \( e(w) \). In terms of the weights \( q_i \), the mass in this probability distribution will be shifted out to the higher-order spells. The result will be a new steady state wage distribution that exhibits both a higher mean wage and, in general, increased dispersion.\(^\text{10}\) Thus increases in inequality of the cross-sectional distribution are consistent with welfare-improving changes in the search environment.

A second potential use of the steady state wage distribution is as a solution to the initial conditions problem. If one is willing to assume that the data are generated by transitions consistent with the process when it has reached the steady state, the steady state wage distribution can be used as a basis for forming the distribution of point-sampled wages. This elegant solution is not particularly useful in this application because the cohorts we follow are very young during

\(^\text{10}\) Increases in the rate of arrivals parameter \( \lambda_e \) have effects on the steady state wage distribution other than a shift in the "mixing" probabilities \( q \). Increases in \( \lambda_e \) increase the value of unemployed search, but by less than they increase the value of being employed at any wage \( w(e) \). As a result, the reservation wage \( w^* \) decreases, which implies that the population wage offer distribution \( F_1 \) and all other order statistic distributions \( F_j, j = 2, \ldots \) will shift as well. These new order statistic distributions may exhibit more or less dispersion than the original ones, which makes it impossible to derive an unambiguous comparative statics result for all parent distributions \( F \).
the sample period and since such a high proportion of sample cases have missing job spell order information. Given the particularly low rate of movement between states observed in the Lombardian data, the time necessary to reach the steady state would extend well beyond the end of our observation period. Because of these concerns we have not used the steady state distribution for estimation purposes in this study.

Our solution has been an agnostic one, which has the advantage of producing consistent estimators under a number of possible violations of modelling assumptions at the cost of discarding a large amount of sample information. We have chosen to condition on the wage associated with any job in progress at the beginning of the sample period. While this method ignores a substantial amount of sample information, we found that most of the maximum likelihood estimates produced using it were reasonably precise.

4.2. Accounting for job-to-job shifts involving wage losses

The strict model of wage-maximizing on-the-job search implies that any direct job-to-job transition should be associated with a wage gain. As we saw in Table 1, while this is in fact what is observed in the majority of sample cases, it is by no means true for all such movements. For the maximum likelihood estimator to be well-defined, some adjustment to the model must be made to get rid of these probability-0 events.

We adopt the method most often employed in the literature, which is to allow for measurement error in wages of some simple type (see, e.g. Wolpin (1987) and van den Berg and Ridder (1996)). The assumption made here is that

\[ \tilde{w} = w + \varepsilon, \]

where \( w \) indicates the “true” wage rate associated with a spell, \( \tilde{w} \) is the observed wage, and \( \varepsilon \) is a random variable which is independently and identically log-normally distributed, so that the density of \( \varepsilon \) is

\[ m(\varepsilon) = \phi((\ln(\varepsilon) - \mu_\varepsilon)/\sigma_\varepsilon)/(\sigma_\varepsilon), \quad \varepsilon > 0, \]

where the parameter \( \sigma_\varepsilon > 0 \) and \( \phi \) is the standard normal probability density function. We restrict the parameters \( \mu_\varepsilon \) and \( \sigma_\varepsilon \) so that

\[ E(\tilde{w}|w) = w \]
\[ \Rightarrow E(\varepsilon|w) = 1 \forall w \]
\[ \Rightarrow \sigma_\varepsilon = (-2\mu_\varepsilon)^{0.5}. \]

where the last line follows from the fact that the mean of \( \varepsilon \) is equal to \( \exp(\mu_\varepsilon + 0.5\sigma_\varepsilon^2) \). This condition thus places an implicit restriction on \( \mu_\varepsilon \), namely, that it must be negative.

The advantage of this approach is that it is simple to implement, and the case for there being measurement error in survey reports on wages is compelling (see, e.g. Duncan and Hill, 1985). The disadvantages are that, especially within a highly nonlinear estimation problem, functional form assumptions regarding the distributional properties of the measurement error are required and are ultimately arbitrary.

4.3. Individual heterogeneity

Individuals within each labour market environment are potentially differentiated in terms of observable and unobservable (to the analyst) characteristics. The introduction of observable heterogeneity, in this case completed schooling levels, is straightforward. The behavioural model is simply estimated for each of the distinct schooling classes that have been defined. No
restrictions have been placed on primitive parameters across different schooling groups, so each group (within each country) is treated as inhabiting its own labour market.

There are many important characteristics that differentiate individuals that are unobservable to the analyst. In principle it is possible to incorporate such factors, as was done in Eckstein and Wolpin (1995), through the use of a finite mixture model. In its simplest implementation, the existence of 2 types of individuals is postulated with each type allowed to have their own vector of primitive parameters $\Omega_i$, $i = 1, 2$, and with the probability that an individual is of type 1 given by $\pi \in (0, 1)$. We attempted to fit such a model using the total sample for each country (i.e. not disaggregating by schooling level). While we were successful in the case of the U.S., we could not obtain estimates of this model using the data from Lombardia. The reason for the failure was that the estimator attempted to force the sample members who were unemployed over the entire 17-month period into one type. Given a type of agent that is continually unemployed, say type 1, the parameter vector $\Omega_1$ is not identified since “permanent” unemployment can be produced by $\lambda_n = 0$, $\eta = \infty$, $b = \infty$, or any combination of these conditions. Thus the lack of identification of the finite mixture model for Lombardia is a result of the short observation period and the relatively long unemployment spells experienced by Italian labour market participants. While these identification issues could have been overcome if we were willing to restrict elements of $\Omega_1$ and $\Omega_2$, the imposition of such restrictions would have been exceedingly arbitrary. In the end, we opted for the inclusion of only an observable characteristic (schooling) in the analysis.

4.4. The maximum likelihood estimator

Under the recoverability condition defined by Flinn and Heckman (1982), all of the parameters of the on-the-job search model are identified with the exception of the discount rate $\rho$ and the cost of search $b$ (which are not jointly identified). If $\rho$ is fixed, all remaining parameters are then identified in principle.

For purposes of writing down the likelihood in a succinct fashion, we follow Wolpin (1992) and break the data for each individual into components we will refer to as “cycles.” A cycle begins with a nonemployment spell. Cycles are defined over the observation period as follows. If an individual begins the sample period in the nonemployment state, he remains in that cycle until such time as he leaves the initial spell of nonemployment and enters a new one. Since every nonemployment spell ends in employment, the individual remains in the same cycle through all successive jobs he holds after the initial nonemployment spell (given that all job-to-job transitions are direct). If an individual begins the observation period in a job, he remains in the original cycle until such time as he experiences a nonemployment spell. In theory an individual can experience an indefinitely large number of cycles over any finite-length sample period. We will let $C_i$ denote the number of cycles experienced by sample member $i$ over the sample period.

For reasons of computational tractability we will utilize duration and wage information only from the first two jobs in each cycle. This will not affect the consistency properties of the estimators, though throwing away information will lead to an efficiency loss. The likelihood function will incorporate information pertaining to whether a second job spell completed during the observation period is followed by another job spell or nonemployment.

11. As shown in Flinn and Heckman (1982), the Pareto distribution does not satisfy what they term the “recoverability condition.” This is not a problem in this application since we have assumed that the lower truncation point of the wage offer distribution is equal to the reservation wage, thus eliminating this potential indeterminacy.
In defining the likelihood function we will utilize the following notation (the individual subscript i has been dropped for notational simplicity):

- $X_{n,c} = 1$ if there is a nonemployment spell in cycle c
- $X_{1,c} = 1$ if there is a first job in cycle c
- $X_{2,c} = 1$ if there is a second job in cycle c
- $X_{3,c} = 1$ if there is a third job in cycle c
- $t_{n,c}$ duration of nonemployment spell in cycle c
- $t_{1,c}$ duration of first job in cycle c
- $t_{2,c}$ duration of second job in cycle c
- $w_{1,c}$ observed wage in first job in cycle c
- $w_{2,c}$ observed wage in second job in cycle c
- $r_{n,c} = 1$ if the nonemployment spell in cycle c is censored
- $r_{1,c} = 1$ if the first job spell in cycle c is censored
- $r_{2,c} = 1$ if the second job spell in cycle c is censored.

Whenever a variable is undefined in a particular cycle (for example, if there is no second job spell in the cycle then $t_{2,c}$, $w_{2,c}$, and $r_{2,c}$ are undefined) we set it equal to zero by convention. Note that the indicator variables for the presence of right-censoring need only be defined for the last cycle in the observation period ($C_i$ for each i).

In the absence of initial conditions problems, the likelihood contribution for a given individual can be written as:

$$l = \prod_{c=1}^{C} \int \int w^* \int w_1 \left[ \frac{1}{n} \exp(-h_{n}t_{n,c}) \right]^{X_{n,c}} \times \left\{ \exp\left(-D(w_1)t_{1,c}\right) \left[ (\lambda_e \tilde{F}(w_1))^{X_{1,c}} \eta^{1-X_{2,c}} \right]^{1-r_{1,c}} \left[ m(\tilde{w}_{1,c}/w_1) / w_1 \right] \right\}^{X_{1,c}} \times \left\{ \exp\left(-D(w_2)t_{2,c}\right) \left[ (\lambda_e \tilde{F}(w_2))^{X_{3,c}} \eta^{1-X_{3,c}} \right]^{1-r_{2,c}} \left[ m(\tilde{w}_{2,c}/w_2) / w_2 \right] \right\}^{X_{2,c}} \times \frac{f(w_2)}{f(w_1)} \frac{f(\tilde{w}_1)}{\tilde{F}(w_1)} \frac{1}{\tilde{F}(w_2)} \,dw_2d\tilde{w}_1, \tag{15}$$

where $D(w) \equiv \eta + \lambda_e \tilde{F}(w)$ and $h_{n} \equiv \lambda_n F(w^*)$. Note that $m(\tilde{w}_{j,c}/w_j) / w_j$ is the density of the observed wage in the jth job in the cycle under the measurement error specification (11) and (12); the term $w_j^{-1}$ is the Jacobian of the transformation.

As was pointed out above, we do not have the information available which would allow us to use (15) to estimate model parameters because of the initial conditions problem, i.e. missing job order information. We have elected to simply condition on the wage in any employment spell in progress at the beginning of our observation period. However, we do not observe the true wage in this case, but only have access to a noisy reading on it. Under our measurement error assumptions, we do know that the true wage is related to the observed wage as follows:

$$w = \tilde{w} / \varepsilon. \tag{16}$$

Then the density of the “true” wage in the sampled job spell is

$$m(\tilde{w}_s/w_s) \tilde{w}_s / w_s^2 \Gamma(\tilde{w}_s), \quad w_s > w^*, \tag{17}$$

where the $s$ subscript signifies sampled spell, the term $\tilde{w}_s / w_s^2$ is the Jacobian of the transformation, and $\Gamma(\tilde{w}_s)$ is a normalizing constant which ensures that the density integrates to unity.

Given the true wage in the sampled spell, the distribution of the wage in any immediately successive spell is $f(w') / \tilde{F}(w_s)$, $w' > w_s$. Conditional on the true wage associated with any
sampled employment spell the duration distribution of the sampled employment spell from the sampling time until the completion of the spell \([t]\) is given by

\[
D(w_s) \exp(-D(w_s)t),
\]

which is identical to the population density of conditional (on the wage rate) job spell durations since the population distribution is exponential. If the sampled job spell is not completed by the end of the observation period, the probability of this event conditional on the wage is given by \(\exp(-D(w_s)T)\), where \(T\) is the length of the observation period. For a sampled nonemployment spell which is completed before \(T\) the likelihood contribution is \(h_n \exp(-h_nT)\) and if such a spell is incomplete the contribution is \(\exp(-h_nT)\). Due to the stationarity of the model, the densities and survivor functions associated with the forward recurrence times of the sampled spells are exactly the same as their population counterparts. Thus the only change in the likelihood function (15) which is required is the substitution of the sampled wage density (17) for the population density of the wage associated with the first job in the first cycle in the observation period when the first cycle begins with a job spell.

In terms of the parameterization of the likelihood function, we have directly estimated the set \(\{\lambda_n, \lambda_e, w^*, \alpha(F), \eta, \mu\}\). Thus the reservation wage characterizing the decision rule of nonemployed individuals is treated as a parameter. As was done in Flinn and Heckman (1982), given estimates of this set of parameters and an assumption concerning the instantaneous discount factor \(\rho\), the point estimate of \(b\) is found as follows. For notational simplicity define

\[
A(w) = \int_{w}^{\infty} V_e(x) dF(x)
\]

for \(w > w^*\). Now we can write the value of search as

\[
V_n = \frac{b + \lambda_n A(w^*)}{\rho + \lambda_n F(w^*)},
\]

or alternatively we can express it as

\[
V_e(w^*) = V_n
\]

\[
\Rightarrow V_n = \frac{w^* + \eta V_n + \lambda_e A(w^*)}{\rho + \eta + \lambda_e F(w^*)}
\]

\[
\Rightarrow V_n = \frac{w^* + \lambda_e A(w^*)}{\rho + \lambda_e F(w^*)}.
\]

Note that in (20), given the reservation wage the value of nonemployment is not a function of \(b\). We can substitute (20) into (2) to get

\[
V_e(w) = \frac{w + \eta \left[ \frac{w^* + \lambda_e A(w^*)}{\rho + \lambda_e F(w^*)} \right] + \lambda_e A(w)}{\rho + \eta + \lambda_e F(w)}. \quad (21)
\]

Call the right-hand side (RHS) of (21) \(S(V_e, w^*)\). Then we can solve the integral equation \(V_e = S(V_e, w^*)\) using m.l. point estimates of all required structural parameters, the m.l. estimate of \(w^*\), and the assumed value of \(\rho\); call the solution of the integral equation \(\hat{V}_e\). Finally, we can solve for \(b\) using

\[
b = \left[ \frac{\rho + \lambda_n F(w^*)}{\rho + \lambda_e F(w^*)} \right] (w^* + \lambda_e A(w^*)) - \lambda_n A(w^*). \quad (22)
\]

The continuity of the RHS of (22) insures that the invariance property of maximum likelihood estimators holds. Thus the estimate of \(b\) obtained by substituting m.l. point estimates into (22), including the estimated function \(V_e\), is consistent and has a \(\sqrt{N}\) asymptotic normal distribution.

12. It should be noted that the model estimated in Flinn and Heckman made no allowance for on-the-job search (that is, \(\lambda_e = 0\)), so that the explicit expression for \(b\) was substantially different.
In all of the empirical work conducted we have assumed that the offer distribution is Pareto. As was shown in Flinn and Heckman (1982), with access only to data containing accepted wage offers and the duration of time spent in labour market states, parametric assumptions are required to estimate a structural search model. We have also estimated the model under the assumptions that the population wage offer distribution was (i) exponential; (ii) log normal; and (iii) half-normal (i.e. a mean 0 normal distribution truncated from below at 0). While model inferences can change dramatically depending on the distributional assumption being employed, this is not the case if all wage offer distributions are truncated at the reservation wage $w^*$. Since we take this as a condition which should always be imposed to guarantee a reasonable level of robustness with respect to distributional assumptions on $F$, there is little to choose among the distributional assumptions in the end. I have chosen to work with the Pareto for its simplicity. Evidence presented below suggests that the fit of the model to certain features of the data is reasonable under this functional form assumption. Nonetheless, the reader should bear in mind that inferences drawn from the model are likely to be at least somewhat sensitive to assumptions regarding the functional form of the wage offer distribution.

Table 3 contains estimates of the search model for the ISTAT sample, both for the aggregate and when disaggregated by schooling class. The estimates of the rates of receiving offers in the nonemployed and employed states and the rates of dismissal provide a good summary of the purely descriptive evidence presented in Table 1. Since the temporal unit of measurement is the month, and since by construction all wage offers are acceptable (because the smallest wage offer received is at least as large as the reservation wage), $\hat{\lambda}_n = 0.079$ implies that unemployment spells in Lombardia last a little over a year on average (i.e. 1/0.079 months). The estimate of the job arrival rate while employed, $\hat{\lambda}_e = 0.007$, implies that a worker in Lombardia must wait almost 12 years before getting a job offer from another firm on average. The good news is that, given $\hat{\alpha} = 0.0021$, a worker in Lombardia can expect to wait almost 40 years before being dismissed or laid-off from their job.

Looking at the model estimates by schooling group, the broad patterns found for the aggregate continue to hold. This is particularly true for the offer and dismissal rate parameters, where there are no (pairwise) significant differences in the estimates of $\lambda_n$, $\lambda_e$, and $\eta$ across the groups. While the null hypothesis of no difference in the entire parameter vector across the three schooling groups is decisively rejected using a likelihood ratio test, this is mainly attributable to the differences in the estimates of the parameters determining the wage offer distribution, $w^*$ and $\alpha$, and of the measurement error parameter. As we might expect, the reservation wage estimates are increasing in the “quality” of the educational credential. To make it easier to compare the wage offer distributions, at the bottom of each column we present the implied mean and squared coefficient of variation of the wage offer distribution ($C_\mu^2(w)$). As we expect, the mean wage offer is increasing in the level of educational credentials, but perhaps more striking is the pattern of inequality in the wage offer distribution. For the first two schooling groups there is very little dispersion in the wage offer distribution, while there is evidence of a very large degree of inequality in wage offers to the group of individuals with a college degree (laurea). In fact, the squared coefficient of variation associated with the wage offer distribution

---

13. For example, under the log normal distributional assumption, where the distribution is truncated from below at the minimum wage, the log likelihood value is $-8549.779$ for the aggregate U.S. data and is $-516.278$ for the aggregate Lombardian data. These values are very close to those obtained under the Pareto assumption and the estimates of all Poisson parameters are essentially constant across the two specifications.

14. One of the reasons for this result is the low number of transitions that occur in the Lombardian sample. Disaggregation by schooling group exacerbates the problem of obtaining precise estimates of these parameters given the small sample size.
of university graduates is more than an order of magnitude greater than what is found for those with lower levels of schooling.

Turning our attention to the first column of Table 4, we note that (for the aggregate samples) the rate of receiving offers when unemployed is about three times greater in the U.S. than in Lombardia, and the rate of receiving job offers when employed is over ten times greater. On the other hand, the rate of dismissals is over an order of magnitude larger in the U.S. sample. These estimates are very much in accord with conventional wisdom regarding the functioning of the markets in the two countries. The Italian market is often thought of as one in which individuals spend long periods of time searching for their first job, but once they find it, tend to keep it for a long period of time, often their entire working life. The U.S. labour market is thought to be one with rapid rates of turnover between jobs, both for "voluntary" and "involuntary" reasons. All of these gross generalizations are consistent with the evidence presented in the first columns of Tables 3 and 4.

In terms of the results for each of the schooling groups in the NLSY sample, we note that more statistically significant differences in arrival rate parameters emerge. The rate of receiving offers when unemployed is increasing in years of schooling until we reach the college graduates (S4), who are receiving job offers at an appreciably lower rate than those with some college (S3). Conversely, we see that the rate of receiving offers when employed is generally declining in years of schooling completed (the estimates of \( \lambda_e \) for the S2 and S3 groups are essentially identical). The estimates imply that employed college graduates have to wait almost 17 months between job offers in comparison with high school dropouts (S1) who wait about one year. The advantage that the high school dropouts have with respect to the offer arrival rate when employed is balanced by their very large dismissal rate. Average time to dismissal, assuming no intervening offers of alternative employment, is 2-6 years for high school dropouts, as compared with 6-9, 4 and 6 years for those with high school diplomas, some college and a college degree, respectively.
The reservation wage is seen to be increasing in the schooling completion level except for the group with some college (S3), though the estimated value of \( w^* \) for this group is not significantly different from that for the high school graduates. The differences in the estimates of the parameter \( \alpha \) are not large given the sizes of the standard errors. Given the large sample size and the noticeable differences in some of the estimated parameters across the schooling groups, it is not surprising that the null hypothesis of equality of the parameter vectors is decisively rejected using a likelihood ratio test.

Monetary quantities are not directly comparable across the two samples, but due to our choice of monetary units the wage offer distributions in the two environments appear to be relatively similar. The mean and standard deviation of the population wage offer distribution are both slightly greater in the NLSY (aggregate) sample. Across schooling groups in the NLSY sample, the mean wage offer is increasing in schooling level, with the exception of the group S3 whose members have some college education. The squared coefficient of variation of the wage offer distribution is approximately equal for those in the two lowest schooling classes, and is highest for the group with some college education. \( C^2_f(w) \) is only about twice as large for the highest schooling group than for the lowest, which is marked contrast to what we found in the sample from Lombardia.

Before conducting the inequality analysis using the model estimates, we present some evidence concerning the ability of the model to fit the data. While formal testing of model fit is possible in principle, because effective sample sizes are small and most of the duration information comes from censored spells, such tests are likely to be inconclusive. This is particularly the case with respect to model estimates for specific schooling groups. Instead, some informal evidence for the aggregate Lombardian and U.S. samples regarding model fit is presented in Figure 4. Figures 4(a) and (c) plot histograms of the wages obtained on the first job in an employment spell along with the implied density of these same wages from the model. The implied density from the model is the density of the “observed” wage offers, \( \bar{w} = w e \), where

<table>
<thead>
<tr>
<th>Parameter</th>
<th>All</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \lambda_n )</td>
<td>0.219</td>
<td>0.184</td>
<td>0.237</td>
<td>0.262</td>
<td>0.214</td>
</tr>
<tr>
<td>(0.010)</td>
<td>(0.015)</td>
<td>(0.021)</td>
<td>(0.023)</td>
<td>(0.020)</td>
<td></td>
</tr>
<tr>
<td>( \lambda_x )</td>
<td>0.082</td>
<td>0.104</td>
<td>0.093</td>
<td>0.097</td>
<td>0.060</td>
</tr>
<tr>
<td>(0.006)</td>
<td>(0.015)</td>
<td>(0.016)</td>
<td>(0.013)</td>
<td>(0.008)</td>
<td></td>
</tr>
<tr>
<td>( w^* )</td>
<td>6.190</td>
<td>6.201</td>
<td>6.392</td>
<td>5.551</td>
<td>7.641</td>
</tr>
<tr>
<td>(0.222)</td>
<td>(0.360)</td>
<td>(0.461)</td>
<td>(0.352)</td>
<td>(0.510)</td>
<td></td>
</tr>
<tr>
<td>( \alpha )</td>
<td>4.169</td>
<td>5.343</td>
<td>5.899</td>
<td>3.747</td>
<td>4.155</td>
</tr>
<tr>
<td>(0.370)</td>
<td>(0.982)</td>
<td>(1.166)</td>
<td>(0.539)</td>
<td>(0.683)</td>
<td></td>
</tr>
<tr>
<td>( \eta )</td>
<td>0.018</td>
<td>0.032</td>
<td>0.012</td>
<td>0.021</td>
<td>0.014</td>
</tr>
<tr>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>( b )</td>
<td>0.262</td>
<td>4.566</td>
<td>1.585</td>
<td>-0.800</td>
<td>-3.082</td>
</tr>
<tr>
<td>(3.129)</td>
<td>(3.562)</td>
<td>(8.177)</td>
<td>(4.739)</td>
<td>(8.763)</td>
<td></td>
</tr>
<tr>
<td>( \mu )</td>
<td>-0.108</td>
<td>-0.107</td>
<td>-0.124</td>
<td>-0.090</td>
<td>-0.107</td>
</tr>
<tr>
<td>(0.007)</td>
<td>(0.012)</td>
<td>(0.015)</td>
<td>(0.012)</td>
<td>(0.013)</td>
<td></td>
</tr>
<tr>
<td>( E_F(w) )</td>
<td>8.143</td>
<td>7.629</td>
<td>7.785</td>
<td>7.571</td>
<td>10.063</td>
</tr>
<tr>
<td>( C^2_f(w) )</td>
<td>0.111</td>
<td>0.056</td>
<td>0.050</td>
<td>0.153</td>
<td>0.112</td>
</tr>
<tr>
<td>( L )</td>
<td>-8550.273</td>
<td>-2059.625</td>
<td>-2338.190</td>
<td>-2017.252</td>
<td>-2071.713</td>
</tr>
<tr>
<td>( N )</td>
<td>1516</td>
<td>264</td>
<td>496</td>
<td>321</td>
<td>435</td>
</tr>
</tbody>
</table>
$w$ is generated by the Pareto distribution and $\varepsilon$ is drawn from the log normal distribution. In the case of Lombardia, the fit is exceptionally good, but we should bear in mind the fact that the histogram is based on only 48 wage observations. For the total U.S. sample, there is somewhat less correspondence between the implied density and the sample histogram, although many of the discrepancies might be attributable to the clustering of the 248 wage reports at focal points and the smaller amount of smoothing used in constructing the histogram. These qualifications aside, it appears that the correspondence between the model prediction and the observed wage distribution is reasonably good in both cases.

Figures 4(b) and (d) contain plots of the hazard rates implied by the model and observed in the data. The hazard rate out of the unemployment state is constant in our stationary model, with the implied monthly exit rate being 0.219 in the U.S. and 0.079 in Lombardia. We have constructed the estimates of the unemployment hazard rates from the data as follows. Since we do not know the beginning date of the unemployment spell when the respondent begins the sample period in this state, all left-censored unemployment spells were ignored. The unemployment spell information used to estimate the hazard function is taken from all spells that begin after the first week of January 1988, whether or not they were completed by the end of the sample period (May 1989). What Kalbfleisch and Prentice (1980) refer to as the “life table method” is used to estimate the hazard rates from this sample of right-censored and non-right-censored unemployment durations. This estimator is essentially a nonparametric maximum likelihood estimator, given the assumption that hazard rates are constant over monthly intervals. This estimator imposes little or no structure on the distribution of unemployment spell lengths, at the cost of requiring large numbers of completed spells for precise estimation. Unfortunately, we only have 14 spells available from the Lombardian data, seven of which are complete. The situation for the NLSY
data is significantly better, where we have access to 355 completed spells of unemployment out of a total of 465. Needless to say, the hazard rate functions are much more precisely estimated for the U.S., though in both cases hazard rates at high duration levels are much less precisely estimated than at low duration levels as a result of monotonically declining sample sizes.

The hazard rate estimates from the Lombardian data show some indication of positive duration dependence, though the extremely small sample size makes any interpretation of these results problematic. The line in Figure 4(b) at the value of 0.079 corresponds to the value predicted by the model. Since this hazard function estimate uses data from all schooling groups, the composition effects induced by mixing should lead us to find evidence of negative duration dependence if the stationary model is the correct one. Thus the fact that positive duration dependence is observed is somewhat disturbing, but given the sample size no conclusion regarding model fit can reasonably be drawn.

The nonparametric hazard function estimates using the aggregate NLSY sample tell a different story. Here there is a clear indication of negative duration dependence, but, as mentioned above, such a pattern is consistent with constant hazards at a more disaggregated level. Having observed substantial differences in the estimates of \( \lambda_n \) for the four schooling groups, the observed pattern could be produced by mixing over these four groups with the constant hazard assumption holding within each. Our conclusion is that there appears to be some slippage between the stationarity assumption and the hazard function estimates. However, a considerable degree of the discrepancy can be accounted for by the heterogeneity in arrival rate parameters for the schooling groups. The stationary analysis we have conducted should not be too misleading, particularly when using the disaggregated (by schooling group) data.

We now turn to a consideration of the implications of these estimates for dispersion in lifetime welfare outcomes. The Italian labour market has been characterized as one in which welfare outcomes are “compressed” relative to what is experienced by participants in more competitive markets like that of the U.S. This view is commonly supported by appealing to cross-sectional evidence on wages or earnings like that presented in Figures 1(a) and (b). As we hope has been made clear, such evidence is very difficult to interpret in a life-cycle setting. Our simulation exercise utilizes point estimates of the search models to compute welfare levels associated with a large number of pseudo-labour market careers.

The simulation exercise itself is straightforward (it is described in detail in Appendix B). The procedure begins by fixing values for all the structural parameters of the model at the m.l. point estimates. Each “individual” enters the labour market in the nonemployment state—then a random number is generated which determines the duration of time he spends in the state until finding a job. The wage associated with the first job is determined by another random number draw and the population distribution of (first job) accepted wage offers. Other random numbers are drawn to determine when new offers are received, their values, and the times of exogenous terminations. The process is repeated until the labour market career has lasted at least 45 years (540 months). For each labour market environment one million sample histories were created. Each sample history is given a single value by integrating the discounted values of all labour market states over the 45 year career. The values of the inequality measures reported

15. In results not presented here, nonparametric hazard function estimates clearly display negative duration dependence for the group of individuals who did not complete high school. There is also an indication of some form of negative duration dependence for those with a high school diploma, though the pattern is less regular. There is no indication of anything other than a constant hazard of leaving unemployment for the two groups with higher educational qualifications.

16. For the computations reported in Table 5 we have actually used all of the m.l. point estimates except for \( \hat{b} \), which has been set to 0. This parameter was very imprecisely estimated, and our desire was not to allow this parameter to unduly influence the inequality comparisons. The results obtained when using the estimated \( \hat{b} \) are not substantially different (see Table A.1)
Table 5 contains the results of the inequality decompositions of the lifetime welfare distributions. The structure of the model induces a dynamic form of averaging that results in significantly lower amounts of dispersion in the lifetime welfare distribution than exist in the cross-sectional wage distribution. In particular, very high wage draws are eventually lost as a result of exogenous separation. Very low, but acceptable, wage draws are lost either due to exogenous separation or as a result of encountering a higher wage offer. In this manner both lower and upper tails of the wage distribution are compressed. The extent to which this process operates depends most directly on the rates of transition between labour market states and the discount rate.

Table 5 contains the results of the inequality decompositions of the lifetime welfare distributions. The structure of the model induces a dynamic form of averaging that results in significantly lower amounts of dispersion in the lifetime welfare distribution than exist in the cross-sectional wage distribution. In particular, very high wage draws are eventually lost as a result of exogenous separation. Very low, but acceptable, wage draws are lost either due to exogenous separation or as a result of encountering a higher wage offer. In this manner both lower and upper tails of the wage distribution are compressed. The extent to which this process operates depends most directly on the rates of transition between labour market states and the discount rate.

In the case of Lombardia, where transition rates are relatively low, the aggregate lifetime welfare distribution is only about 50% as unequal as the cross-sectional wage distribution in terms of the squared coefficient of variation measure, and is even less disperse using the two entropy measures. Even more striking than these differences are those seen within the various educational groups. For the two groups of individuals with lower levels of schooling (S1 and S2), the life cycle inequality is approximately an order of magnitude lower than cross-

for the lifetime welfare distribution and the steady state wage distribution are computed from the sample distributions of these labour market career values. It is important to note that the measurement error distribution is not considered in any of these calculations, that is, the wage draws are taken from the distribution of \( w \) and not the distribution of \( \tilde{w} = w \varepsilon \). This is consistent with our interpretation of \( \varepsilon \) as representing pure measurement error unrelated to the utility yield of the labour market state.

Table 5 contains the results of the inequality decompositions of the lifetime welfare distributions. The structure of the model induces a dynamic form of averaging that results in significantly lower amounts of dispersion in the lifetime welfare distribution than exist in the cross-sectional wage distribution. In particular, very high wage draws are eventually lost as a result of exogenous separation. Very low, but acceptable, wage draws are lost either due to exogenous separation or as a result of encountering a higher wage offer. In this manner both lower and upper tails of the wage distribution are compressed. The extent to which this process operates depends most directly on the rates of transition between labour market states and the discount rate.

In the case of Lombardia, where transition rates are relatively low, the aggregate lifetime welfare distribution is only about 50% as unequal as the cross-sectional wage distribution in terms of the squared coefficient of variation measure, and is even less disperse using the two entropy measures. Even more striking than these differences are those seen within the various educational groups. For the two groups of individuals with lower levels of schooling (S1 and S2), the life cycle inequality is approximately an order of magnitude lower than cross-

<table>
<thead>
<tr>
<th>Group</th>
<th>n_s/n</th>
<th>( \bar{v} )</th>
<th>( C^2 )</th>
<th>T</th>
<th>L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lombardia (ISTAT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>2781.499</td>
<td>0.056</td>
<td>0.019</td>
<td>0.015</td>
</tr>
<tr>
<td>Within class:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S1</td>
<td>0.614</td>
<td>2636.617</td>
<td>0.009</td>
<td>0.004</td>
<td>0.004</td>
</tr>
<tr>
<td>S2</td>
<td>0.335</td>
<td>2758.904</td>
<td>0.006</td>
<td>0.003</td>
<td>0.003</td>
</tr>
<tr>
<td>S3</td>
<td>0.051</td>
<td>4674.173</td>
<td>0.166</td>
<td>0.055</td>
<td>0.044</td>
</tr>
<tr>
<td>Weighted sum within:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.033</td>
<td>0.008</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>Between:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.023</td>
<td>0.011</td>
<td>0.009</td>
<td></td>
</tr>
<tr>
<td>Proportion between (IW):</td>
<td></td>
<td>0.411</td>
<td>0.579</td>
<td>0.600</td>
<td></td>
</tr>
</tbody>
</table>

| U.S. (NLSY79) |       |                 |          |       |       |
| Total         | 1     | 3502.810        | 0.037    | 0.017 | 0.017 |
| Within class: |       |                 |          |       |       |
| S1            | 0.134 | 2798.808        | 0.009    | 0.004 | 0.004 |
| S2            | 0.357 | 3308.814        | 0.010    | 0.005 | 0.005 |
| S3            | 0.206 | 3221.473        | 0.020    | 0.009 | 0.008 |
| S4            | 0.304 | 4220.064        | 0.020    | 0.009 | 0.008 |
| Weighted sum within: |     |                 |          |       |       |
|                |       | 0.016           | 0.007    | 0.007 |       |
| Between:       |       |                 |          |       |       |
|                |       | 0.021           | 0.010    | 0.010 |       |
| Proportion between (IW): |   | 0.568           | 0.588    | 0.588 |       |

17. The inequality measures for the steady state wage distribution are computed as follows. We let the simulation process run for 10,000 months for each labour market participant, and we compute the inequality measure from the distribution of wages in the 10,000th month, which should be a more than adequate approximation to the steady state distribution in both environments.
sectional inequality. In contrast, the $C^2$ of the life cycle distribution for the college-trained (S3) is slightly greater than it is for the cross-sectional distribution of wages for this group. In terms of the other inequality measures, the amount of inequality is the same or slightly less for the career value distribution for those with a laurea. Because of the large differences in inequality in lifetime welfare across the schooling groups in Lombardia, the “between” variation in lifetime inequality is four or five times higher than was the case for the cross-sectional wage distributions.

For the U.S. data, the declines in inequality when moving from the cross-sectional to the lifetime welfare distributions are even more dramatic for the aggregate sample. Across all three of the inequality measures the declines are approximately one order of magnitude. Comparing the upper and lower panels of Table 5, we see that the amount of life cycle inequality in the two labour markets is almost the same. Recall that there is approximately three times as much inequality in the U.S. when comparing cross-sectional wage distributions.

The amount of inequality within schooling groups in the U.S. declines more than an order of magnitude when using the life cycle welfare distributions. Life cycle welfare inequality is found to be about twice as large for the two groups with more schooling than for those groups with no college experience. After “transitory” welfare differences are smoothed away, the amount of between variation in inequality increases to about 58% of total inequality. This is approximately the same as was found in the case of Lombardia.

Let us take a moment to summarize our findings using estimates from the complete samples for each country or simulations based on those point estimates. In terms of the cross-sectional wage distributions observed in January 1988, the amount of inequality in the U.S. sample was approximately three times greater. Using our model estimates we found that differences in inequality in the wage offer distributions were much less dramatic, with $C^2$ equal to 0.071 in Lombardia and 0.110 in the U.S. Differences in the steady state wage distributions were more pronounced, with $C^2$ for Lombardia equal to 0.198 compared with a value of 0.371 for the U.S. Since in our two samples average labour market experience is approximately 10 years, we also computed the implied distribution of wages for individuals with 10 years of experience. For these distributions $C^2$ is similar to what was found in the cross-sectional wage distributions, taking the value 0.125 for Lombardia and 0.483 for the U.S.\footnote{18} Finally, we found that our model estimates implied similar levels of inequality in the distributions of the value of the labour market career for Lombardia and the U.S.

Given that we found rough similarity in the wage offer distributions in the two countries, it is worth emphasizing that differences in other theoretical “cross-sectional” wage distributions, namely the steady state wage distributions and wage distributions observed for a cohort with 10 years of labour market experience, are produced by differences in the spell-order distribution. These spell-order distributions appear in Table 6. We note that the steady state spell-order distributions are very similar in the two countries. While this may appear surprising, it results from the fact that both the rate of arrival of job offers when employed and the exogenous separation rate are approximately an order of magnitude less in Lombardia than in the U.S. On the other hand, there are very large differences in the job-spell distributions observed after 10 years of experience, and, as a result, in the corresponding wage distributions. These differences arise from the large discrepancy in the rates of convergence to the steady state in these two markets. While 73% of those in the Lombardian market are found in the first job of an employment spell after 10 years of experience, 62% of those in the U.S. market are in a second or higher order job spell if employed.

\footnote{18. We cannot make too much of the ability of the model to “fit” the cross-sectional wage distribution in terms of $C^2$ however, since measurement error should be added to the structural distribution for us to make a valid comparison.}
TABLE 6

<table>
<thead>
<tr>
<th>Job spell</th>
<th>Distribution</th>
<th>&quot;Ten year&quot; Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lombardia</td>
<td>U.S.</td>
</tr>
<tr>
<td>1</td>
<td>0.594</td>
<td>0.549</td>
</tr>
<tr>
<td>2</td>
<td>0.274</td>
<td>0.283</td>
</tr>
<tr>
<td>3</td>
<td>0.103</td>
<td>0.125</td>
</tr>
<tr>
<td>4</td>
<td>0.025</td>
<td>0.037</td>
</tr>
<tr>
<td>5</td>
<td>0.004</td>
<td>0.006</td>
</tr>
</tbody>
</table>

FIGURE 5

(a) Proportion employed Lombardia; (b) proportion employed U.S.; (c) growth in mean wages Lombardia; (d) growth in mean wages U.S.; (e) growth in S.D. of wages Lombardia; (f) growth in S.D. of wages U.S.

Some of these results are quite striking, and it is natural to question whether the model which generates them has any claim to be taken seriously. We can use our simulated labour market histories to partially address this question. In particular, we look at the manner in which the mean and standard deviation of wages change over the life cycle. Even though the model is highly stylized, and neglects such admittedly important phenomena as within job wage change and business cycle effects on the parameters characterizing the labour market environment, Figure 5 provides some evidence that the model predictions are broadly consistent with evidence obtained using panel data to estimate flexible models of earnings dynamics.

Life cycle phenomena are produced in a stationary environment as the process moves from initial conditions to the steady state distribution of events. In our case, all labour market
participants are assumed to share the same initial condition, which is nonemployment. The length of time it takes for the distribution of state occupancies to be well-approximated by the steady state distribution depends on the transition rates which characterize the process. Moreover, the distributions of state occupancies as indexed by labour market age of the cohort will move in a systematic fashion since all individuals begin their careers unemployed. We expect to see the mean wage of the employed monotonically increase as a function of labour market age; predictions regarding trends in dispersion seem to be dependent on the functional form of the wage offer distribution.

In Figure 5(a), we see that for Lombardia the model predicts that unemployment will be heavily concentrated in the early years of labour market participation, which is consistent with substantial amounts of empirical evidence for Italy. After the first few years in the labour market, however, the Italian employment rate is predicted to be a few points higher than that of the U.S., which is graphed in Figure 5(b). Because the rate of receiving offers is so much greater in the U.S., and because all offers are accepted under our modelling assumptions, the steady-state employment rate is achieved within a few years in the U.S., in comparison with 5 or 6 for Italy. There is no pronounced concentration of unemployment in the early years of labour market participation for the U.S. 19

Figures 5(c) and (d) contain plots of mean wage growth using the U.S. and Lombardian estimates. We see that growth is always positive in the two environments, though the rate of growth is dramatically greater in the U.S. over the first 8 to 10 years of labour market participation. The growth rate in Italy is higher in the later years of life, simply because the U.S. wage distribution has already converged to its steady state value (in which there is no growth due to our assumption of a time-invariant wage offer distribution), while the Italian distribution will not converge to the steady state distribution until long after a normal working lifetime is completed. Of course, this life-cycle growth in wages is built into our model, so as such it is not a surprising finding. What is interesting is the predicted relative growth rates in mean earnings for the two countries.

Figures 5(e) and (f) complete our brief look at life cycle movements with plots of the standard deviation of the wage distribution as a function of labour market age. We see that in both countries wage inequality grows monotonically as the population ages. This observation is consistent with findings from many panel data studies of earnings dynamics. The rate of convergence to the steady state value of the standard deviation is similar to what was observed in the case of the mean values.

6. CONCLUSION

We provide evidence that the cautionary tale told by Friedman is well worth heeding when comparing levels of inequality in very different institutional environments. Labour market environments which exhibit less inequality in the cross-section may produce no less, or even more, inequality in the long run. We showed that this was exactly the case with respect to a comparison between a very large region in Italy and the U.S., using samples composed of young male labour market participants. We showed, however, the rough equivalence in life-cycle inequality in the two markets was mainly attributable to the extremely large amounts of inequality found in the welfare distributions of the highly educated in Lombardia. Inequality among individuals with lower levels of educational credentials was found to be much less in Lombardia than in the U.S.

19. Recall that these data pertain to white males only. We would expect to see much longer spells of unemployment and more concentration of unemployment among youth if the sample were extended to include other racial groups.
The analysis is based on a standard model of off- and on-the-job search. The search framework serves three purposes. First, it gives us the ability to use the data at our disposal, which is based on a very brief observation period of 17 months, to estimate parameters which can generate outcomes over the entire labour market career of participants. Second, it provides a metric, however limited, for making welfare comparisons. Third, it provides a way to associate characteristics of the labour market with the distribution of welfare outcomes. We would argue that even when data on remuneration are available for a substantial fraction of each sample member’s labour market career, the use of a model (not necessarily one of on-the-job search) greatly facilitates the conversion of a remuneration history into a summary welfare measure and provides a readily interpretable way to map characteristics of the environment into moments of the welfare distribution.

While the evidence we have presented might lead us to label Italy (or its stand-in here, Lombardia) as the status society, there are obviously a number of qualifications to make. As we have remarked previously, the model is very stylized, and is particularly suspect due its omissions of business cycle variability in the search environment, on-the-job wage variability, and self-employment status. Superior data could be used to rectify these first two problems at the cost of considerable increases in model complexity. While the on-the-job search model can be generalized to allow for search in the dependent and independent worker (i.e. self-employment) sectors, the interpretation of search model parameters in the self-employment sector is problematic. Because our initial attempts to estimate such a generalized model using the Lombardian data produced results consistent with what is reported in this paper, we have chosen to ignore the self-employed. Nonetheless, this may be a serious omission simply because the self-employment state is much more common among Italian sample members than among those from the U.S. It is not clear how the neglect of all of these phenomena in both environments would tend to bias our comparisons of welfare inequality, but clearly our inferences could easily change.

A more direct measurement issue concerns the reporting of labour market activity in the two environments. It is well known that there is a significant “underground” economy in Italy. Because this unofficial labour market is illegal, little is known about its exact size or nature, though we can assume that there is nothing of analogous importance in the U.S. (in particular for the demographic group of which our sample is composed). We cannot be certain as to whether respondents to a government-financed survey would be reluctant to report holding an “illegal” job, though it is reasonable to suppose that they might. If so, nonreporting of such jobs might seriously affect our estimates of transition rates between labour market states in Lombardia, with resulting repercussions for our comparisons of welfare inequality. There seems little of a constructive nature to do about this problem beyond being aware of its existence.

In addition, we believe that a complete analysis of the issues addressed here requires a careful consideration of differences in capital markets in the two environments. While capital markets in Lombardia and the U.S. seem to be relatively similar now, there still exist substantial differences in requirements for loans and the types of savings instruments available. Given any realization of a lifetime stream of earnings, different capital market environments can produce grossly different lifetime welfare outcomes. The incorporation of such considerations into this type of comparative analysis is an important research goal.

Despite all of these qualifications, we hope that at a minimum the analysis has provided a clear cautionary note for the interpretation of differences in cross-sectional remuneration measures across economic environments. While other dynamic modelling frameworks would have assuredly produced different numerical results, this general point is not model-specific nor is its scope limited to the two economic environments analysed here.
APPENDIX A. DERIVATION OF THE STEADY-STATE WAGE DISTRIBUTION

In this section we explicitly derive the mapping between the dynamic behavioural model and the steady state cross-sectional earnings distribution. Before proceeding to the details of the derivation, we present a heuristic explanation of this mapping.

When individuals are sampled at some random point in time, agents who happen to be employed at the sampling time can be distinguished by the number of the job they are currently holding in the on-going employment spell. For example, in Figure 2 individual B was found in his first job in the employment spell while individual A was found to be in his fourth. The individual who has had three previous jobs in the current spell will on average have a higher wage than individual B since he has had the maximum of four draws from the acceptable wage offer distribution while individual B has had only one draw.

To compute the steady state wage distribution requires that we first find the wage distribution by job order, and then find the steady state distribution of job orders in on-going employment spells. The first task is relatively straightforward, being essentially an exercise involving order statistics. To compute the steady state job order distribution, we first compute the steady state distribution of the number of jobs in an employment spell. Conditional on there being $k$ jobs in a completed employment spell, we then derive the probability that an individual will be sampled while in his $j$th job, $j = 1, \ldots, k$. Then the probability that an individual will be found in his first job in an employment spell will be the steady state probability of having a one-job employment spell plus the conditional probability of being in the first job when sampled randomly given the employment spell ends after two jobs multiplied by the probability that the employment spell ends after two jobs, etc. We now proceed to a formal derivation of these distributions.

By the structure of the model all nonemployed individuals accept any draw from the distribution $F$ which exceeds the reservation wage associated with the nonemployment state, $w^*$. Then the density of wages in the first job in an employment spell is

$$f_1(w_1) = \frac{f(w_1)}{\hat{F}(w^*)} \chi[w_1 \geq w^*],$$

where $\chi[Z]$ takes the value $1$ if $Z$ is true and equals $0$ if $Z$ is false.

To derive the marginal density of the wages associated with the second job in an employment spell, we begin by writing the joint density of first and second job wages,

$$f_{1,2}(w_1, w_2) = f_{2|1}(w_2 | w_1) f_1(w_1)$$

$$= \frac{f(w_2)}{\hat{F}(w_1)} \chi[w_2 > w_1] \frac{f(w_1)}{\hat{F}(w^*)} \chi[w_1 > w^*]$$

$$= \frac{f(w_2)}{\hat{F}(w^*)} \frac{f(w_1)}{\hat{F}(w_1)} \chi[w_2 > w_1 > w^*],$$

where the form of the conditional density $f_{2|1}$ follows from the fact that the second job is a random draw from $f$ truncated from below at the first job wage $w_1$. Integrating over $w_1$ to get the marginal density of $w_2$ we have

$$f_2(w_2) = \frac{f(w_2)}{\hat{F}(w^*)} \int_{w^*}^{w_2} h(w_1) dw_1 \chi[w_2 > w^*],$$

$$= \frac{f(w_2)}{\hat{F}(w^*)} \hat{F}_2(w_2) \chi[w_2 > w^*],$$

where $\hat{F}_2(w_2)$ (for "adjustment factor for the second job wage density evaluated at $w_2$") is the term in brackets in (27) and $h(x)$ is the hazard function associated with the distribution $F$ evaluated at $x$, or $h(x) = f(x)/\hat{F}(x)$.

By induction we can construct the marginal density of wages associated with the $k$th job in an employment spell as

$$f_k(w_k) = \frac{f(w_k)}{\hat{F}(w^*)} \hat{F}_k(w_k) \chi[w_k > w^*],$$

where the adjustment factor $\hat{F}_k(w_k)$ is defined recursively by

$$\hat{F}_k(w_k) = \int_{w^*}^{w_k} \hat{F}_{k-1}(x) h(x) dx.$$
that the employment spell ends after the first job is the likelihood that the first job ends due to a dismissal, which is \( \frac{\eta}{(\eta + \lambda_e \hat{F}(w_1))} \). Then the unconditional probability of an employment spell ending after one job is

\[
R(1) = \int w^* \frac{\eta}{\eta + \lambda_e \hat{F}(w_1)} f_1(w_1) dw_1.
\]

(31)

Conditional on not having exited an employment spell before the \( k \)th job, we have by extension that the probability that the employment spell ends after the \( k \)th job (which we will denote \( \hat{R}(k) \)) is

\[
\hat{R}(k) = \int w^* \frac{\eta}{\eta + \lambda_e \hat{F}(w_k)} f_k(w_k) dw_k, \quad k = 2, 3, \ldots.
\]

(32)

Then the unconditional distribution of numbers of jobs in completed employment spells is

\[
R(k) = \prod_{j=1}^{k-1} (1 - \hat{R}(j)) \hat{R}(k), \quad k = 2, 3, \ldots.
\]

(33)

with \( R(1) \) defined in (31) and \( \hat{R}(1) = R(1) \).

The second step in the process involves the computation of the conditional probabilities of being found in job \( j \) when sampled given that the employment spell contains \( k \geq j \) jobs. Denote this generic conditional probability by \( \pi_{jk} \), and define the matrix \( \mathcal{P} \) [which is a square upper triangular matrix of countably infinite dimension] by

\[
\mathcal{P} = \begin{bmatrix}
1 & \pi_{12} & \pi_{13} & \cdots \\
0 & \pi_{22} & \pi_{23} & \cdots \\
0 & 0 & \pi_{33} & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}.
\]

(34)

The fact that \( \pi_{11} = 1 \) is definitional.

Consider the determination of the elements in the second column of \( \mathcal{P} \). Now the total time spent in employment spells which end after two jobs (denoted \( t^{(2)} \)) is the sum of the first job spell duration conditional on the first job ending in a quit and the second job spell duration conditional on the second job ending in a dismissal. Conditional on the wages associated with these two jobs, the spell durations are independently distributed. The conditional expectation of the duration of a two-job employment spell is then

\[
E_t^{(2)}(w_1, w_2) = \int_0^\infty \int_0^\infty (t_1 + t_2) \lambda_e \hat{F}(w_1) \exp(-D(w_1)t_1) \eta \exp(-D(w_2)t_2) dt_1 dt_2
\]

\[
= \int_0^\infty t_1 \lambda_e \hat{F}(w_1) \exp(-D(w_1)t_1) dt_1 + \int_0^\infty t_2 \eta \exp(-D(w_2)t_2) dt_2
\]

\[
= \lambda_e \hat{F}(w_1)D(w_1)^{-2} + \eta D(w_2)^{-2}.
\]

(35)

(36)

(37)

where \( D(w) \) is defined as \( \lambda_e \hat{F}(w) + \eta \). Then the unconditional expectation of the duration of a two-job spell is

\[
E_t^{(2)} = \int_{w^*} \int_{w^*} \left[ \lambda_e \hat{F}(w_1)D(w_1)^{-2} + \eta D(w_2)^{-2} \right] f_{1,2}(w_1, w_2) dw_1 dw_2
\]

\[
= \int_{w^*} \lambda_e \hat{F}(w_1)D(w_1)^{-2} f_1(w_1) dw_1 + \int_{w^*} \eta D(w_2)^{-2} f_2(w_2) dw_2.
\]

(38)

(39)

To simplify notation, define

\[
\mathcal{E}_k = \int_{w^*} \lambda_e \hat{F}(w_k)D(w_k)^{-2} f_k(w_k) dw_k
\]

(40)

and

\[
\mathcal{W}_k = \int_{w^*} \eta D(w_k)^{-2} f_k(w_k) dw_k.
\]

(41)

Using standard ergodic arguments, the probability of finding an individual in the first job of a two-job spell under random sampling is

\[
\pi_{12} = \frac{\mathcal{E}_1}{\mathcal{E}_1 + \mathcal{W}_2}
\]

and of course \( \pi_{22} = 1 - \pi_{12} \). In general, the expected duration of an employment spell which ends after the \( k \) jobs is

\[
E_t^{(k)} = \mathcal{E}_1 + \mathcal{E}_2 + \cdots + \mathcal{E}_{k-1} + \mathcal{W}_k
\]

\[
= T_{k-1} + \mathcal{W}_k.
\]

(42)

(43)

(44)
where $T_{k-1} \equiv \sum_{i=1}^{k-1} \xi_i$. Since the probability of finding an individual in job spell $j$ of an employment spell which contains a total of $k$ jobs is equal to the average duration of job spell $j$ in a $k$ job employment spell divided by the expected duration of a $k$ job employment spell, we have

$$\pi_{jk} = \begin{cases} \frac{\xi_j}{T_{k-1} + W_k} & \text{if } j = 1, \ldots, k-1; \quad k = 2, 3, \ldots \\ \frac{W_k}{T_{k-1} + W_k} & \text{if } j = k; \quad k = 2, 3, \ldots \\ 0 & \text{if } j > k; \quad k = 1, 2, \ldots. \end{cases}$$ (45)

Now we can define the probability distribution for finding an individual in the $j$th job of an employment spell given that they are employed at the random sampling time as

$$Q = \mathcal{P}R.$$ (46)

**APPENDIX B. SIMULATION METHODS**

To compute the welfare values $\omega(1), \ldots, \omega(N)$ corresponding to a given set of values for the structural parameters, the procedure used was as follows. All "hypothetical" careers begin in the unemployment state. The discount rate was set to $\rho/12$ (recall that this is a monthly rate), where we set $\rho = 0.04$. The generic spell is indexed by $i$, and it begins at time $r_i$ ($r_1 = 0$). The labour market career is terminated at the conclusion of the first spell with a termination date after the 540th month of the career, which corresponds to the individual spending approximately 45 years in the labour market.

The instantaneous remuneration rate attached to spell $i$ is denoted $r_i$, and the type of spell is denoted $d_i$, where $d_i = 1$ if spell $i$ is an employment spell and is equal to 0 if it is an unemployment spell, and the total duration of the spell is denoted $t_i$.

If spell $i$ was an unemployment spell (so that $d_i = 0$), we first generated a draw $t_i$ from an exponential distribution with parameter $\lambda d \hat{F}(\omega^*)$. We then generated a wage draw $w_{i+1}$ (since this wage is associated with the $i + 1$st spell) from the accepted wage offer distribution $F(w | w \geq \omega^*)$.20 The contribution of spell $i$ to lifetime welfare is given by

$$V_i = \exp(-\rho r_i) \int_0^{t_i} b \exp(-\rho u) du$$ (47)

$$= \rho^{-1} \exp(-\rho r_i) [1 - \exp(-\rho t_i)] b.$$ (48)

The next (employment) spell would then begin at calendar time $r_{i+1} = r_i + t_i$ at the wage $w_{i+1}$.

If spell $i$ is an employment spell the procedure is essentially the same except for the fact that there are two ways to exit the spell, either through an exogenous separation or a quit into a higher-paying job. Let the wage associated with the job be given by $w_i$. We first took a draw from an exponential distribution with parameter $\eta d \hat{F}(\omega^*)$, which gave us the duration of the employment spell, $t_i$. We then generated a draw $x$ from a uniform distribution on the unit interval. If $x < \eta \hat{F}(\omega_i)$ the spell was considered to have ended in a dismissal so that spell $i + 1$ is a nonemployment spell; conversely, if $x > \eta \hat{F}(\omega_i)$ the next spell was an employment spell. When spell $i + 1$ was an employment spell, a wage rate was generated from the distribution $F(w | w > \omega_i)$. Given that spell $i$ is an employment spell, its contribution to lifetime welfare is

$$V_i = \rho^{-1} \exp(-\rho r_i) [1 - \exp(-\rho t_i)] w_i.$$ (49)

Let $N$ denote the number of spells which commenced prior to the 540th month. Say that this history has been generated for "individual" $j$. Then individual $j$'s labour market career has value

$$\omega(j) = \sum_{i=1}^{N} V_i$$ (50)

under the particular set of structural parameter values utilized in the experiment.

Acknowledgements. This research was partially supported by the C.V. Starr Center for Applied Economics at New York University. I am especially grateful to Ugo Trivellato and Enrico Rettore for making the Italian data used in this research available to me. Helpful comments were received from seminar participants at Bocconi, Padova, Penn, the IRP at the University of Wisconsin, DELTA, the Equilibrium Wage Distributions Conference (Paris), the Savings Workshop (Tilburg), the European University Institute, and NYU. Orazio Attanasio, Giuseppe Bertola, Tito Boeri, Meta Brown, 20. Under our assumption that the lower bound of the support of the Pareto is equal to the reservation wage, this distribution is the same as $\hat{F}$.
TABLE A.1

<table>
<thead>
<tr>
<th>Group</th>
<th>$n_s/n$</th>
<th>$\bar{V}$</th>
<th>$C^2$</th>
<th>$T$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lombardia (ISTAT)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>2774.089</td>
<td>0.055</td>
<td>0.018</td>
<td>0.014</td>
</tr>
<tr>
<td>Within class:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S1$</td>
<td>0.614</td>
<td>2617.883</td>
<td>0.010</td>
<td>0.005</td>
<td>0.004</td>
</tr>
<tr>
<td>$S2$</td>
<td>0.335</td>
<td>2832.146</td>
<td>0.004</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>$S3$</td>
<td>0.051</td>
<td>4273.331</td>
<td>0.259</td>
<td>0.087</td>
<td>0.072</td>
</tr>
<tr>
<td>Weighted sum within:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion between (IW):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>U.S. (NLSY79)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>3523.327</td>
<td>0.031</td>
<td>0.014</td>
<td>0.013</td>
</tr>
<tr>
<td>Within class:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S1$</td>
<td>0.134</td>
<td>3078.912</td>
<td>0.005</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>$S2$</td>
<td>0.357</td>
<td>3346.562</td>
<td>0.011</td>
<td>0.005</td>
<td>0.005</td>
</tr>
<tr>
<td>$S3$</td>
<td>0.206</td>
<td>3204.348</td>
<td>0.021</td>
<td>0.009</td>
<td>0.009</td>
</tr>
<tr>
<td>$S4$</td>
<td>0.304</td>
<td>4142.954</td>
<td>0.023</td>
<td>0.010</td>
<td>0.009</td>
</tr>
<tr>
<td>Weighted sum within:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion between (IW):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

REFERENCES

ANAND, S. (1983), Inequality and Poverty in Malaysia (New York: Oxford University Press (for the World Bank)).


