9 Employment and the Labor Force

9.1 Introduction

In this chapter the employment and labor force sector will be discussed. The employment part of the sector is based on the work in Fair [19], where the short-run demand for workers and for hours paid-for per worker was examined in considerable detail. The labor force part of the sector is less sophisticated and consists essentially of two rather simple labor force participation equations. In Section 9.2 the employment equation will be developed and estimated, and in Section 9.3 the labor force equations will be discussed. The chapter concludes in Section 9.4 with a summary of the sector and with a discussion of how the sector is treated within the context of the overall model. Much of the discussions in Sections 9.2 and 9.3 follows closely the discussion in Fair [18].

In order of causality in the model, the price sector should actually be discussed before the employment and labor force sector, since real output feeds into the employment and labor force sector from the price sector. The price sector, however, uses the labor force equations (though not the labor force predictions) in the development of a potential GNP series, and it is thus more convenient to discuss the employment and labor force sector first.

9.2 The Short-Run Demand for Employment

In macroeconomic models the link between output changes and employment changes is generally provided either through an aggregate production function or an aggregate employment demand function. If an employment demand function is used, it is frequently derived from a production function. It was argued rather extensively in Fair [19] that any attempt to estimate the parameters of a short-run production function in the standard way is doomed to failure, because the true labor inputs are not observed. A critical distinction was made in [19] between the (observed) number of hours paid-for per worker and the (unobserved) number of hours actually worked per worker, and it was argued that the latter is not likely to be equal to the former except during peak output periods. Using this distinction, a model of short-run employment demand was developed in [19] and was estimated for a number of three-digit manufacturing industries. The concept of "excess labor" played an important role in the model, and the estimated amount of excess labor on hand for an industry appeared to be a significant determinant of the change in employment for that industry. The present study extends the model developed in [19] to the total private nonfarm sector. It will be seen that the estimated amount of excess labor on hand in the private nonfarm sector does appear to be a significant determinant of the change in employment in the sector. The following is a brief outline of the model.

The Concept of Excess Labor

Let M_t denote the number of workers employed during period t, HP_t the average number of hours paid-for per worker during period t, H_t the average number of hours actually worked per worker during period t, and H_t^* the standard number of hours of work per worker during period t. If HP_t is greater than H_t, then firms are paying workers for more hours than they are actually working, i.e., firms are paying for "nonproductive" hours. (HP_t can never be smaller than H_t, since hours worked must be paid for.) If output and the short-run production function are taken to be exogeneous, then the two variables at the firm's command in the short run are M_t and HP_t.[1] If the total number of man hours paid-for, $M_t HP_t$, is greater than total number of man hours worked, $M_t H_t$, the firm can decrease either M_t or HP_t or both.

In the present model the desired distribution of $M_t HP_t$ between M_t and HP_t is assumed to be a function of H_t^*. H_t^* is the dividing line between standard hours of work and more costly overtime hours: if HP_t is greater than H_t^*, then an overtime premium has to be paid on the hours above H_t^*. It is thus assumed that the long-run equilibrium number of hours paid-for per worker is H_t^*. With this in mind, the measure of excess labor is taken to be $\log H_t^* - \log H_t$, which is the (logarithmic) difference between the standard number of hours of work per worker and the actual number of hours worked per worker.[2] If H_t is less than H_t^*, there is considered to be a positive amount of excess labor on hand (i.e., too few hours worked per worker and thus too

[1] From the short-run production function below, once output and M_t are determined, H_t is automatically determined.

[2] For reasons that will be clearer below, the functional form of the model is taken to be the log-linear form. In order to ease matters of exposition and where no ambiguity is involved, in what follows the difference of the logs of two variables (e.g., $\log H_t^* - \log H_t$) will be referred to merely as the difference of the variables.

many workers on hand), and if H_t is greater than H_t^*, there is considered to be a negative amount of excess labor on hand (i.e., too many hours worked per worker and thus too few workers on hand).[3] How the amount of excess labor on hand is assumed to affect changes in employment will be discussed below.

The Short-Run Production Function

The production function inputs are taken to be the number of man hours worked and the number of machine hours used. The short-run production function is assumed to be characterized by (1) no short-run substitution possibilities between workers and machines and (2) constant short-run returns to scale both with respect to changes in the number of workers and machines used and with respect to changes in the number of hours worked per worker and machine per period. Let Y_t denote the amount of output produced during period t, let M_t continue to denote the number of workers employed during period t, and let K_t denote the number of machines used during period t. Under the assumption that there are no completely idle workers or machines (which will be made here), assumption (1) implies that the number of hours worked per worker, H_t, is equal to the number of hours worked per machine. This is discussed in more detail in Fair [19], Chapter 3, but basically all it says is that if a fixed number of workers is required per machine, then it is not possible to have workers and machines working a different number of hours.

Assumptions (1) and (2) imply that the short-run production function is

$$Y_t = \min\{\alpha_t M_t H_t, \beta_t K_t H_t\}, \tag{9.1}$$

where α_t and β_t are coefficients that may be changing through time as a result of technical progress. The assumption that there are no completely idle workers or machines implies that $\alpha_t M_t H_t$ equals $\beta_t K_t H_t$ in (9.1), so that (9.1) implies

$$Y_t = \alpha_t M_t H_t. \tag{9.2}$$

Equation (9.2) has been taken to be the basic production function in this study.[4]

---

[3] In some industries a certain amount of overtime work has become standard practice—workers expect it and firms are reluctant not to grant it—and for these industries H_t^* should be considered to be the standard number of hours of work per worker plus this standard or "accepted" number of overtime hours of work per worker. In other words, H_t^* should be considered to be the *desired* number of hours paid-for and worked per worker.

[4] See Fair [19], Chapter 3, for a more complete discussion of the derivation of this production function.

The Measurement of Excess Labor in the Private Nonfarm Sector

In [19] it was argued that when attempting to estimate the parameters of a production function, seasonally unadjusted data should be used. A production function is a technical relationship between certain physical inputs and a physical output, not a relationship between seasonally adjusted inputs and a seasonally adjusted output. Unfortunately perhaps, the world of empirical macroeconomics is largely a seasonally adjusted world, and much of the national income accounts data are not even published on a seasonally unadjusted basis. Consequently, for the work in this study seasonally adjusted data have been used. Because of this and because of the highly aggregated nature of the data anyway, much less reliance can be put on the conclusions reached in this study than on those reached in the study of three-digit industries in [19]. The present study is merely an attempt to use some of the ideas and conclusions in [19] to develop an aggregate employment equation that can be used for forecasting purposes. It should not be considered to be an attempt to test various hypotheses about short-run employment demand.

The data on Y_t, M_t, and HP_t that have been used to estimate the employment equation below are data for the private nonfarm sector. There appears to be little systematic short-run relationship between output and employment in the agricultural and government sectors: an attempt to explain agricultural and government employment in the same way that private nonfarm employment was explained did not meet with much success. The employment data for the agricultural sector are not very good, however, and the poor results for this sector may have been due in large part to measurement errors.[5] Whatever the reason for the poor results, the decision was made to treat both agricultural and government employment as exogenous in the model. The ability to forecast these variables exogenously will be examined in Chapter 13.

The data on Y_t, M_t, and HP_t are described in Table 9–1. The data on private nonfarm output, Y_t, are national income accounts data and are currently published in the *Survey of Current Business*. The data on private nonfarm employment, M_t, and on hours paid-for per private nonfarm worker, HP_t, are compiled by the Bureau of Labor Statistics (BLS). The data on M_t and HP_t used in this study were obtained directly from the BLS, but some of the data are currently published in index number form in the *Monthly Labor Review*, Table 32. The data on M_t and HP_t are designed by the BLS

[5] See [38], pp. 123–129 for a discussion of the lack of quality of much of the agricultural data.

Table 9–1. List and Description of the Variables Used in the Employment and Labor Force Sector.

For the Employment Equation

M_t	= Private Nonfarm Employment in thousands of workers, SA (primarily establishment data).
HP_t	= Hours Paid-for per Private Nonfarm Worker in hours per week per worker, SA.
Y_t	= Private Nonfarm Output in billions of 1958 dollars, SA, annual rates.

For the Labor Force Equations

MA_t	= Agricultural Employment in thousands of workers, SA (primarily household survey data).
MCG_t	= Civilian Government Employment in thousands of workers, SA (primarily establishment data).
E_t	= Total Civilian Employment in thousands of workers, SA (household survey data).
D_t	= Difference between the establishment employment data and the household-survey employment data in thousands of workers, SA ($= M_t + MA_t + MCG_t - E_t$).
AF_t	= Level of the Armed Forces in thousands.
LF_{1t}	= Level of the Primary Labor Force (males 25–54) in thousands, SA (household survey data).
LF_{2t}	= Level of the Secondary Labor Force (all others over 16) in thousands, SA (household survey data).
P_{1t}	= Noninstitutional Population of males 25–54 in thousands.
P_{2t}	= Noninstitutional Population of all others over 16 in thousands.
UR_t	= The Civilian Unemployment Rate, SA (household survey data) ($= 1 - E_t/[LF_{1t} + LF_{2t} - AF_t]$).

Note: SA = Seasonally Adjusted.

to cover the same sector of the economy as the sector covered by the national income accounts data on private nonfarm output. Note that it is hours paid-for per worker that are observed (HP_t) and not necessarily hours actually worked per worker (H_t). The data on M_t and HP_t that have been used in this study are presented in Appendix A. The data on Y_t can be easily obtained from the *Survey of Current Business*.

Using the data on Y_t, M_t, and HP_t, excess labor in the private nonfarm sector is measured as follows. In Figure 9–1 output per paid-for man hour, Y_t/M_tHP_t, is plotted for the 471–694 period. The dotted lines in the figure are peak-to-peak interpolation lines of the series. The assumption is made that at each of the interpolation peaks Y_t/M_tHP_t equals Y_t/M_tH_t, i.e., that output per paid-for man hour equals output per worked man hour. From equation (9.2) this provides an estimate of α_t at each of the peaks. The further assumption is then made that α_t moves smoothly through time along the interpolation lines from peak to peak. This assumption provides estimates of α_t for each quarter of the sample period, which from (9.2) and from the data on Y_t allows an estimate of man-hour requirements, M_tH_t, to be made for

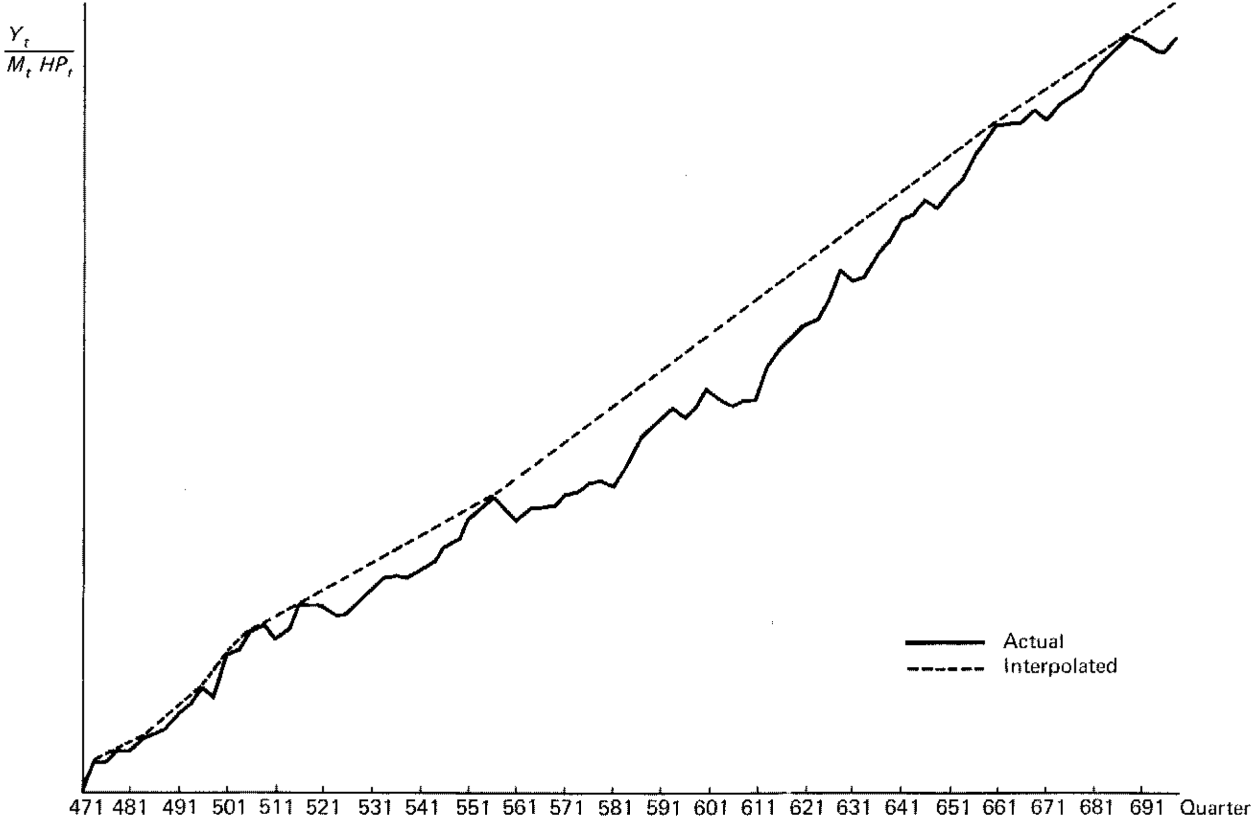

Figure 9.1. Output Per Paid-for Man Hour

each quarter.[6] For any quarter, $M_t H_t$ is the estimated number of man hours required to produce Y_t. If $M_t H_t$ is divided by H_t^*, the standard (or desired) number of hours of work per worker, the result, denoted as M_t^d, can be considered to be the desired number of workers employed for quarter t:

$$M_t^d = \frac{M_t H_t}{H_t^*}. \tag{9.3}$$

M_t^d is the desired number of workers employed in the sense that if man-hour requirements were to remain at the level $M_t H_t$, M_t^d can be considered to be the number of workers firms would want to employ in the long run. In the long run each worker would then be working the desired number of hours.

The amount of (positive or negative) excess labor on hand is then taken to be $\log M_t - \log M_t^d$, which is the (logarithmic) difference between the actual number of workers employed and the desired number. It is easy to show that this measure of excess labor is the same as $\log H_t^* - \log H_t$, which is the measure defined in Section 9.2:

$$\begin{aligned} \log M_t - \log M_t^d &= \log M_t - \log M_t H_t + \log H_t^* \qquad \text{[using (9.3)]} \\ &= \log M_t - \log M_t - \log H_t + \log H_t^* \\ &= \log H_t^* - \log H_t. \end{aligned} \tag{9.4}$$

In other words, the amount of excess labor on hand can be looked upon either as the difference between the number of workers employed and the desired number employed or as the difference between the standard number of hours of work per worker and the actual number of hours worked per worker.

Except for the measurement of H_t^*, the measurement of excess labor on hand in the private nonfarm sector is complete. The production function parameter α_t has been estimated from peak-to-peak interpolations of the output per paid-for man-hour series in Figure 9–1, and from the estimates of α_t and the data on Y_t, measurements of man-hour requirements have been made using the production function (9.2). Using (9.3), man-hour requirements can then be divided by some measure of the standard number of hours worked per worker to yield a series on the desired number of workers employed. The assumption that has been made about the standard number of hours of work per worker will be discussed in the next section.

[6] The 661–684 line was extrapolated to get the 691, 692, 693, and 694 values for α_t. The choice of the peaks in Figure 9–1 is, of course, somewhat arbitrary, although the results were not very sensitive to the choice of slightly different peaks. The 601 and 624 "peaks" were not used as interpolation peaks because demand was still relatively weak during these periods and it seemed likely that output per paid-for man hour was still below output per worked man hour during 601 and 624.

The Short-Run Demand for Workers

In [19], using monthly data at the three-digit industry level, the change in the number of workers employed was seen to be a function of the amount of excess labor on hand and of expected future changes in output of up to six months in advance. Past changes in output were also seen to be significant for a few industries. It was argued in [19] (pp. 51–52), that the past change-in-output variables may help depict the reaction of firms to the amount of excess labor on hand. With respect to future output expectations, it is unlikely that the influence of output expectations more than one quarter ahead can be picked up with the highly aggregative data used in this study. The basic equation explaining employment demand is thus taken to be:

$$\log M_t - \log M_{t-1} = a_1(\log M_{t-1} - \log M_{t-1}^d) + b_0(\log Y_{t-1} - \log Y_{t-2}) + b_1(\log Y_t^e - \log Y_{t-1}). \tag{9.5}$$

Y_t^e is the expected amount of output produced during quarter t. Equation (9.5) states that the change in the number of workers employed during quarter t is a function of the amount of excess labor on hand in quarter $t-1$, of the change in output during quarter $t-1$, and of the expected change in output for quarter t. a_1 is expected to be negative and b_0 and b_1 to be positive. A more complete discussion of the theoretical model upon which equation (9.5) is based is presented in Chapter 3 of [19].

Since M_t is actually the average number of workers employed during quarter t and Y_t the average rate of output during quarter t and since employment decisions are likely to be made on less than a quarterly basis, it will be assumed here that $Y_t^e = Y_t$. In other words, it is assumed that output expectations are perfect for the current quarter.

One more assumption is necessary before equation (9.5) can be estimated. This is the assumption regarding the standard number of hours of work per worker, H^*. It is assumed that H^* is either a constant or a slowly trending variable, and specifically that

$$H_{t-1}^* = \bar{H}e^{qt}, \tag{9.6}$$

where $\bar{H}$ and q are constants. Using this assumption and the definition of M_t^d in (9.3), the excess labor variable in equation (9.5) can then be written

$$\log M_{t-1} - \log M_{t-1}^d = \log M_{t-1} - \log M_{t-1}H_{t-1} + \log \bar{H} + qt. \tag{9.7}$$

Using (9.7) and the assumption about Y_t^e made above, equation (9.5) becomes

$$\log M_t - \log M_{t-1} = a_1 \log \bar{H} + a_1 qt + a_1(\log M_{t-1} - \log M_{t-1}H_{t-1}) + b_0(\log Y_{t-1} - \log Y_{t-1}) + b_1(\log Y_t - \log Y_{t-1}). \tag{9.5'}$$

Equation (9.5′) is now in a form that can be estimated. Data on output and employment are available directly, and data on man-hour requirements, $M_{t-1}H_{t-1}$, were constructed in the manner described in the previous section.[7]

There are perhaps two main differences between equation (9.5) and previous aggregate employment equations. One, of course, is the inclusion of the excess labor variable. This variable is designed to measure the reaction of firms to the amount of too little or too much labor on hand. The second difference is that equation (9.5) does not directly include a capital stock variable. It is instead assumed that there are no short-run substitution possibilities between workers and machines and that the long-run effects of the growth of technical progress on employment (as embodied in, say, new capital stock) are reflected in the movement through time of α_t in (9.2). If α_t is increasing through time, then, other things being equal, M_t^d in (9.3) will be falling, since man-hour requirements, $M_t H_t$, will be falling. The amount of excess labor on hand will thus be increasing. The effects of the growth of technology on employment decisions are thus taken care of by the reaction of firms to the amount of excess labor on hand.

The Results

Equation (9.5′) was estimated for the 561–694 period under the assumption of first order serial correlation of the error terms. Since output is taken to be exogenous in the employment and labor force sectors, the two-stage least squares technique described in Chapter 2 did not have to be used to estimate the equation, and the equation was estimated using the simple Cochrane–Orcutt technique. As was done in the money GNP sector, observations for 593, 594, 601, 644, 651, and 652 were omitted from the sample period because

[7] Note that the $\log M_{t-1} - \log M_{t-1}H_{t-1}$ term in (9.5)′ is equal to $-\log H_{t-1}$. Equation (9.5′) is written the way it is to emphasize that man-hour requirements, $M_{t-1}H_{t-1}$, were estimated directly from the above production function.

of strikes. The results were:

$$\begin{aligned}\log M_t - \log M_{t-1} &= \underset{(3.44)}{-.514} + \underset{(1.57)}{.0000643t} - \underset{(3.41)}{.140}(\log M_{t-1} - \log M_{t-1}H_{t-1})\\ &\quad + \underset{(2.34)}{.121}\ (\log Y_{t-1} - \log Y_{t-2}) + \underset{(6.43)}{.298}(\log Y_t - \log Y_{t-1})\\ \hat{r} &= \underset{(2.52)}{.336} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (9.8)\\ \mathrm{SE} &= .00310\\ R^2 &= .778\\ &\qquad\qquad 50 \text{ observ.}\end{aligned}$$

As in previous chapters, the numbers in parentheses are t-statistics (in absolute values) and $\hat{r}$ is the estimate of the serial correlation coefficient. Since the dependent variable is already in first differenced form, the R-squared was computed taking the dependent variable to be in this form rather than in second differenced form.

All of the estimates in (9.8) are of the expected sign, and all but the estimate of the coefficient of the time trend are significant. The estimate of the coefficient of the excess labor variable is $-.140$, which implies that, other things being equal, 14 percent of the amount of excess labor on hand is removed each quarter. The past output change variable, however, is also picking up some of the effect of the reaction of firms to the amount of excess labor on hand. The estimate of the serial correlation coefficient is rather small at .336, but it is large enough to indicate that there is at least some degree of serial correlation present. This contrasts with the three-digit industry results in [19], which gave very little evidence of serial correlation.

Other equations similar to (9.8) were also estimated. The two-quarter-lagged change in output, $\log Y_{t-2} - \log Y_{t-3}$, was added to the equation, and it was not significant. In an effort to test for the effect of future output expectations on the change in employment, $\log Y_{t+1} - \log Y_t$ was added to (9.8) (under the hypothesis ot perfect expectations), and it likewise was not significant.[8] As expected, the aggregate data here do not appear to be capable of picking up any effect of future output expectations on current employment changes. Equation (9.8) was also estimated with $\log M_{t-1}$ replacing the excess labor variable, $\log M_{t-1} - \log M_{t-1}H_{t-1}$, to see if the excess labor variable is perhaps significant in (9.8) merely because it is of the nature of a lagged dependent variable.[9] The results were quite poor and $\log M_{t-1}$

[8] The equation included only 49 observations, since the 694 observation had to be dropped to allow for the last observation for Y_{t+1}.

[9] See the more complete discussion of this in [19], pp. 72–76.

was not significant by itself. The equation,

$$\log M_t - \log M_{t-1} = a_0 + a_1 t + a_2 \log M_{t-1} + a_3 \log Y_t + a_4 \log Y_{t-1},$$

which is common to many of the previous studies of short-run employment demand, was also estimated, and the results again were worse than those in (9.8).[10] Equation (9.8) was thus chosen as the basic equation determining the change in the number of private nonfarm workers employed.

Before concluding this section, the estimates of the amount of excess labor on hand will be examined in a little more detail. Note that from equation (9.8), estimates of q and $\bar{H}$ are available. (See equation (9.5′).) This means that a series on H_t^* can be constructed from equation (9.6). Using this series and the series on man-hour requirements, $M_t H_t$, constructed above, a series on the desired number of workers employed, M_t^d, can be constructed from equation (9.3). These calculations were made, and in Table 9–2 the actual series on M_t, the constructed series on M_t^d, and the difference in these series, $M_t - M_t^d$, are presented for the 561–694 period. The value of $M_t - M_t^d$ in Table 9–2 for any one quarter indicates what the excess labor situation was like for that quarter. In the last quarter of 1969, for example, there were 922,000 too many workers employed for the amount of output produced. This compares with a range of 2,240,000 too few workers in 661 to 2,722,000 too many workers in 611.

It should finally be noted that the employment model described above provides an explanation of why in the short run "productivity" falls when output falls and rises when output rises. The coefficient of $\log Y_t - \log Y_{t-1}$ in equation (9.8) is less than one (.298 to be exact), and thus when output changes by a certain percentage, employment changes by less than this percentage. Employment is then gradually changed over time to its desired level by the reaction of firms to the amount of excess labor on hand (and to the past change in output). Output per worker will thus be positively correlated with output in the short run. Also, from the results in [18] and [19] it can be seen that the number of hours paid-for per worker (HP_t) changes by a smaller percentage than output does in the short run, and indeed that total man hours paid-for ($M_t HP_t$) changes by a smaller percentage than output does. This means that output per paid-for man hour ($Y_t/M_t HP_t$) will also be positively correlated with output in the short run. Therefore, whether productivity is defined as output per worker or output per paid-for man hour, it follows that productivity and output will be positively correlated in the

[10] The fit was slightly worse: $R^2 = .758$ vs. .778 in (9.8), and serial correlation was much more pronounced: $\hat{r} = .610$ vs. .336 in (9.8). Also, as argued in [19], the equation just estimated has little theoretical justification, especially if it is taken as an equation from which a production function parameter can be derived.

short run. Only gradually will employment and hours paid-for per worker adjust to their desired levels. The process by which this adjustment takes place is described in more detail in Chapter 8 [19].

Table 9–2. Estimated Values for M_t^d.

Quarter	M_t	M_t^d	$M_t - M_t^d$	Quarter	M_t	M_t^d	$M_t - M_t^d$
561	54850	55417	−567	631	57461	57277	184
562	55087	55503	−416	632	57763	57444	319
563	54892	54886	6	633	58175	58186	−11
564	55133	55161	−28	634	58294	58717	−423
571	55340	55276	64	641	58738	59480	−742
572	55514	54899	615	642	59196	59951	−755
573	55437	54862	575	643	59499	60417	−918
574	54872	53510	1362	644	59934	60329	−395
581	53811	51571	2240	651	60464	61370	−906
582	53198	51571	1627	652	61011	61961	−950
583	53543	52659	884	653	61608	62993	−1385
584	54034	53692	342	654	62339	64187	−1848
591	54533	54256	277	661	62923	65163	−2240
592	55427	55493	−66	662	63526	65448	−1922
593	55421	54530	891	663	64182	65617	−1435
594	55607	54711	896	664	64472	65963	−1491
601	56250	55701	549	671	64730	65238	−508
602	56410	55028	1382	672	64762	65197	−435
603	56170	54326	1844	673	64948	65607	−659
604	55892	53556	2336	674	65401	65882	−481
611	55773	53051	2722	681	65835	66503	−668
612	55652	54009	1643	682	66368	67508	−1140
613	55929	54754	1175	683	66621	67813	−1192
614	56517	55554	963	684	67020	68125	−1105
621	56964	56108	856	691	67753	68147	−394
622	57361	56668	693	692	68192	68137	55
623	57384	57160	224	693	68526	68128	398
624	57200	57451	−251	694	68736	67814	922

Note: Figures are in thousands of workers.

9.3 The Labor Force and the Unemployment Rate

The purpose of the labor force equations is to allow predictions of the unemployment rate to be made, given predictions of private nonfarm employment (M_t) from the employment equation. There are three problems involved in going from predictions of M_t to predictions of the unemployment rate. First, M_t excludes agricultural and government workers. Secondly, M_t is based primarily on establishment data, and not on the household survey data, which are used to estimate the size of the labor force and the unemployment rate. A link thus has to be found between the establishment-based data and the household survey data. Finally, predictions of the labor force need to be made in order to allow predictions of the unemployment rate to be made.

With respect to the first problem, as mentioned above, agricultural employment and government employment are taken to be exogenous in the model. MA_t will be used to denote the number of agricultural workers employed, MCG_t the number of civilian government workers employed, and AF_t the number of people in the armed forces. The data on these three variables are described in Table 9–1. Data on AF_t can be obtained from data published currently in *Economic Indicators* by subtracting the figures on the civilian labor force from the figures on the total labor force. The data on MA_t and MCG_t, on the other hand, were obtained directly from the BLS. The data differ slightly from the data on agricultural and government workers that are currently published in *Economic Indicators*. Data on MCG_t, for example, exclude government enterprise workers, whereas the data on government workers in *Economic Indicators* include enterprise workers. For the BLS data used in this study, government enterprise workers are included in M_t, since government enterprise output is counted as private output. Likewise, there are a few discrepancies between the MA_t series and the agricultural employment series published in *Economic Indicators* because of the need by the BLS to match the agricultural employment series to the corresponding agricultural output series in the national income accounts.

With respect to the problem of establishment data versus household survey data, let E_t denote the total number of civilian workers employed according to the household survey. The data on E_t are described in Table 9–1 and are currently published in *Economic Indicators*. The difference D_t is then defined to be

$$D_t = M_t + MA_t + MCG_t - E_t. \tag{9.9}$$

D_t is positive and appears to consist in large part of people who hold more than one job. (The establishment series are on a job number basis and the household survey series are on a person employed basis.)

Given that D_t is composed primarily of people who hold more than one job, one would expect that it would respond to labor market conditions, and this appeared to be true from the results achieved here. D_t was taken to be a function of a time trend and M_t, and the following equation was estimated for the 561–694 period (excluding the strike observations 593, 594, 601, 644, 651, and 652) under the assumption of first order serial correlation of the error terms:

$$\begin{aligned} D_t &= -13014 - 71.10t + .358M_t \\ &\qquad (8.23) \qquad (6.15) \qquad (9.39) \\ \hat{r} &= .600 \\ &\quad (5.30) \\ \text{SE} &= 181.4 \\ R\Delta^2 &= .460 \\ &\qquad 50 \text{ observ.} \end{aligned} \tag{9.10}$$

What equation (9.10) says is that, other things being equal, a change in M_t of say, 1000, leads to a change in E_t of only 642. The difference of 358 is taken up either by moonlighters or by other discrepancies between the establishment data and household survey data.

A number of equations similar to (9.10) were also estimated. Black and Russell [2], for example, have estimated an equation similar to (9.10), with the unemployment rate used in place of M_t. This equation was also estimated for the work here, but it led to poorer results than those in (9.10). Slightly less than 50 percent of the variance of the change in D_t has been explained by equation (9.10) and the estimate of the serial correlation is fairly high, but none of the other equations estimated were an improvement over (9.10) and so (9.10) was chosen as the basic equation determining D_t.

Once M_t is determined, D_t can be determined by equation (9.10), and then taking MA_t and MCG_t as exogenous in equation (9.9), E_t can be determined. Since E_t is used in calculating the unemployment rate, this leaves only the labor force to be determined in order to determine the unemployment rate. There are many special factors that are likely to affect labor force participation rates—some of which have been described by Mincer [37]—and only limited success has so far been achieved in explaining participation rates over time. In this study no attempt has been made to develop an elaborate and refined set of participation rate equations. The labor force has been disaggregated only into primary (males 25–54) and secondary (all others over 16) workers, and the specification of the equations has remained simple. The purpose of the work here is merely to see how useful simple participation rate equations can be in forecasting the unemployment rate.

The labor force participation rate of primary workers does not appear to be sensitive to labor market conditions. None of the variables depicting labor market conditions were significant in the participation rate equations estimated here. In the final equation, therefore, the participation rate of primary workers was taken to be a simple function of time. The equation was estimated for the 561–694 period (excluding the six strike observations) under the assumption of first order serial correlation of the error terms. The results were:

$$\begin{aligned}
\frac{LF_{1t}}{P_{1t}} = & \underset{(658.38)}{.981} \quad - \underset{(8.57)}{.000190t} \\
\hat{\rho} = & \underset{(1.94)}{.265} \\
SE = & .00193 \\
R\Delta^2 = & .447 \\
& 50 \text{ observ.}
\end{aligned} \tag{9.11}$$

LF_{1t} denotes the primary (males 25–54) labor force, and P_{1t} denotes the noninstitutional population of males 25–54. Both variables include people in the armed forces. The data on LF_{1t} and P_{1t} are described in Table 9–1. The data are household survey data and were obtained directly from the BLS. P_{1t} is taken to be exogenous in the model. (The ability to forecast P_{1t} exogenously will be discussed in Chapter 13.) Note that less than half the variance of the change in the participation rate has been explained in equation (9.11). The variance of LF_{1t}/P_{1t} is small enough, however, so that the LF_{1t} series does not pose serious difficulties for short-run forecasting purposes.

The participation rate of secondary workers does appear to be sensitive to labor market conditions, but apparently in no simple way. The coefficients of the equations that were estimated in this study were quite sensitive to the choice of the period of estimation, and in particular the large increase in the participation rate from 1965 through 1969 did not appear to be consistent with past behavior. In the final equation chosen, the participation rate of secondary workers was taken to be a function of time and of the ratio of total employment (including armed forces) to total population 16 and over. The equation was estimated for the 561–694 period (excluding the six strike observations) under the assumption of first order serial correlation of the error terms. The results were:

$$\begin{aligned}
\frac{LF_{2t}}{P_{2t}} = & \underset{(2.69)}{.180} + \underset{(4.97)}{.000523t} + \underset{(3.67)}{.447} \frac{\widehat{E_t + AF_t}}{P_{1t} + P_{2t}} \\
\hat{\rho} = & \underset{(9.32)}{.797} \\
SE = & .00228 \\
R\Delta^2 = & .373 \\
& 50 \text{ observ.}
\end{aligned} \tag{9.12}$$

$$\left[1, t, \frac{LF_{2t-1}}{P_{2t-1}}, \frac{E_{t-1} + AF_{t-1}}{P_{1t-1} + P_{2t-1}}, \frac{AF_t}{P_{1t} + P_{2t}}, \frac{M_t}{P_{1t} + P_{2t}}, \right.$$
$$\left. \frac{M_{t-1}}{P_{1t-1} + P_{2t-1}}, \frac{MCG_t}{P_{1t} + P_{2t}}, \frac{MA_{t-1} + MCG_{t-1}}{P_{1t-1} + P_{2t-1}}\right].$$

LF_{2t} denotes the secondary labor force (including armed forces) and P_{2t} denotes the noninstitutional population (including armed forces) of everyone over 16 except males 25–54. Data on LF_{2t} and P_{2t} are described in Table 9–1. Again, the data are household survey data and were obtained directly from the BLS. Like P_{1t}, P_{2t} is taken to be exogenous in the model.

Equation (9.12) is similar to equations estimated by Tella [42], [43], although here the employment population ratio is taken to include all workers and other individuals over 16 and not just secondary workers and other secondary individuals. Other kinds of participation equations for secondary workers were also estimated, but equation (9.12) appeared to give the best results.

There is one obvious statistical problem in estimating an equation like (9.12), which is due to the fact that LF_{2t} and total civilian employment, E_t, are computed from the same household survey. The household survey is far from being error free, and errors of measurement in the survey are likely to show up in a similar manner in both LF_{2t} and E_t. The coefficient estimate of $(E_t + AF_t)/(P_{1t} + P_{2t})$ in an equation like (9.12) will thus be biased upward unless account is taken of the errors of measurement problem. Because of this problem, equation (9.12) was estimated by the instrumental variable or two-stage least squares technique described in Chapter 2. The normal two-stage least squares technique could not be used because of the assumption of serial correlation of the error terms.

The instruments that were used for $(E_t + AF_t)/(P_{1t} + P_{2t})$ in (9.12) are listed in brackets after the equation. As discussed in Chapter 2, the first four instruments listed are necessary in order to insure consistent estimates. The other instruments are based on equations (9.9) and (9.10). Write equation (9.10) as

$$D_t = a_0 + a_1 t + a_2 M_t + u_t, \tag{9.10$'$}$$

where $u_t = r_0 u_{t-1} + e_t$. (The error term e_t is assumed to have mean zero and constant variance and to be uncorrelated with M_t and with its own past values.) Combining equations (9.9) and (9.10′) and solving for E_t yields

$$\begin{aligned} E_t = & -a_0(1 + r_0) + r_0 a_1 - a_1(1 + r_0)t + (1 - a_2)M_t + MA_t \\ & + MCG_t - r_0 E_{t-1} \\ & + r_0(1 - a_2)M_{t-1} + r_0 MA_{t-1} + r_0 MCG_{t-1} - e_t. \end{aligned} \tag{9.13}$$

Since M_t, M_{t-1}, MCG_t, and MCG_{t-1} in equation (9.13) are not computed from the household survey, they are not likely to be correlated with the measurement error in E_t and thus are good instruments to use. In addition, if the measurement errors themselves are serially uncorrelated (which is assumed here), then even though MA is computed from the household survey, MA_{t-1} in equation (9.13) will not be correlated with the measurement error in E_t and thus can be used as an instrument.

When an equation like (9.13) was estimated by the simple Cochrane–Orcutt technique for the same sample period, the coefficient estimate of $(E_t + AF_t)/(P_{1t} + P_{2t})$ was .608—versus .447 in (9.12)—which seems to indicate that, unless corrected, the measurement error bias is quite large in equations like (9.12).

It will be seen in Chapters 11–13 that equation (9.12) does not give particularly good results. The labor force participation of secondary workers grew quite rapidly during the 1965–1969 period—more rapidly than would seem warranted from the growth in the employment–population ratio—and equation (9.12) is not capable of accounting for all of this growth. The within-sample forecasts are reasonable, since the time trend in the equation can pick up most of the unexplained growth, but the outside-sample forecasts are much worse. As will be examined in Chapter 12, the coefficient estimates in (9.12) are not very stable over time, and the equation does a poor job of extrapolating into a period unless it has been estimated through that period. Mincer [37] makes a very compelling argument that many special factors (such as laws relating to Social Security retirement premiums and minimum wages) are likely to influence participation rates, and in a more complete study these factors should be taken into account. Also, the participation rates of much more disaggregated groups should be analyzed. It is beyond the scope of this study to attempt to do this, and to the extent that the labor force participation of secondary workers continues to change in ways not related to the employment–population ratio, equation (9.12) will continue to be one of the weaker equations of the model.

The employment and labor force sector is now complete. Having determined E_t in the manner described above, and taking P_{1t}, P_{2t}, and AF_t to be exogenous, LF_{1t} and LF_{2t} can be determined from equations (9.11) and (9.12). By definition, the civilian labor force is equal to $LF_{1t} + LF_{2t} - AF_t$, and so the civilian unemployment rate can be determined as:

$$UR_t = 1 - \frac{E_t}{LF_{1t} + LF_{2t} - AF_t}. \tag{9.14}$$

All of the data that have been collected for the work in this section are presented in Appendix A. These include data on MA_t, MCG_t, E_t, AF_t,

LF_{1t}, LF_{2t}, P_{1t}, and P_{2t}. The data are quarterly averages of monthly data. Except for AF_t, P_{1t}, and P_{2t}, the data are seasonally adjusted.

9.4 Summary

All of the variables that are used in the employment and labor force sector are listed in Table 9–1. The sector consists essentially of one production function, four behavioral equations, and two identities; and it will be convenient to list these equations in order of their causality in the sector.

$$\text{(i)} \qquad M_t H_t = \frac{1}{\alpha_t} Y_t \qquad (9.2)$$

$$\begin{aligned} \text{(ii)} \quad \log M_t - \log M_{t-1} = & -.514 + .0000643t \\ & - .140(\log M_{t-1} - \log M_{t-1}H_{t-1}) \\ & + .121(\log Y_{t-1} - \log Y_{t-2}) \\ & + .298(\log Y_t - \log Y_{t-1}), \qquad \hat{r} = .336. \qquad (9.8) \end{aligned}$$

$$\text{(iii)} \qquad D_t = -13014 - 71.10t + .358M_t, \quad \hat{r} = .600. \qquad (9.10)$$

$$\text{(iv)} \qquad E_t = M_t + MA_t + MCG_t - D_t. \qquad (9.9)$$

$$\text{(v)} \qquad \frac{LF_{1t}}{P_{1t}} = .981 - .000190t, \qquad \hat{r} = .265. \qquad (9.11)$$

$$\text{(vi)} \qquad \frac{LF_{2t}}{P_{2t}} = .180 + .000523t + .447 \frac{E_t + AF_t}{P_{1t} + P_{2t}}, \qquad \hat{r} = .797. \qquad (9.12)$$

$$\text{(vii)} \qquad UR_t = 1 - \frac{E_t}{LF_{1t} + LF_{2t} - AF_t}. \qquad (9.14)$$

Private nonfarm output, Y_t, is fed into the employment and labor force sector from the price sector, and then the unemployment rate is determined as follows. First, man-hour requirements are determined from (i), α_t having been estimated in the manner described in Section 9.2. Then, using the man-hour requirement estimates, private nonfarm employment is determined from (ii). The difference between the establishment and household survey data is then determined from (iii), which allows total civilian employment to be determined from the definition (iv). The labor force is then determined from (v) and (vi), and finally the unemployment rate is determined from the definition (vii). Aside from Y_t, the exogenous variables in the section are MA_t, MCG_t, P_{1t}, P_{2t}, and AF_t.

With respect to predictions of the unemployment rate, there is some error cancellation in the model that is worth noting. Positive errors in predicting M_t, for example, will lead, other things being equal, to positive errors in predicting D_t in (iii), which will in turn lead to smaller positive errors in predicting E_t. Likewise, errors in predicting E_t will lead, other things being equal, to errors in the same direction in predicting LF_{2t} in (vi), which in turn lead to smaller errors in predicting the unemployment rate.

The accuracy of the sector as a whole will be examined in Chapters 11–13 within the context of the overall model. The accuracy is also examined in Fair [18], where the actual values of output are used rather than the predicted values from the price sector.