EXCESS LABOR AND AGGREGATE EMPLOYMENT FUNCTIONS

Ray C. Fair

Econometric Research Program

Research Memorandum No. 110

August 1969

The research described in this paper was supported by NSF grant GS-1840 and the computer work by NSF grant GJ-34.

Princeton University

Econometric Research Program

207 Dickinson Hall

Princeton, New Jersey

EXCESS LABOR AND AGGREGATE EMPLOYMENT FUNCTIONS

I. INTRODUCTION

For policy purposes knowledge of the link between output changes and employment changes is of obvious importance. In macro economic models this link is generally provided either through an aggregate production function or an aggregate employment demand function (which is many times derived from a production function).[1] It was argued rather extensively in [3] that the attempt to estimate the parameters of a short-run production function is doomed to failure because, among other things, the true labor inputs are not observed. A critical distinction was made in [3] 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.

This paper extends the model of the short-run demand for workers and for hours paid-for per worker developed in [3] to the whole economy, and from the model an attempt is made to provide a link between aggregate output changes and changes in the unemployment rate. As in [3], the concept of "excess labor" plays an important role in the model, and the estimated amount

[1]See Nerlove [7] for a review of the use of production or employment functions in macro economic models.

of aggregate excess labor on hand appears to be a significant determinant of the change in aggregate employment.

In the next section the model of employment demand developed in [3] is briefly outlined and then modified slightly for use with aggregate data. Measurements of the amount of aggregate excess labor on hand are also derived in this section. In Section III the results of estimating the aggregate employment equation and the aggregate hours paid-for per worker equation are presented and analyzed. In Section IV an attempt is then made to provide a link between the employment predictions of Section III and predictions of the unemployment rate. The equations which were developed to provide this link consist of two labor force participation equations and an equation explaining the difference between establishment based employment and household survey based employment. Finally, in Section V the overall model is simulated, both statically and dynamically, for the 561-692 period, and from these results a judgment can be made as to the potential future accuracy of the model.

II. THE MODEL OF EMPLOYMENT DEMAND[1]

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 HS_t the standard number of hours

---

[1]It should be stressed that the following is only a brief outline of the model developed in [3]. For a defense of the assumptions and specifications of the model, the reader is referred to the more complete discussion in [3].

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 "non-productive" hours. (Theoretically, 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 exogenous, 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_tHP_t, is greater than total number of man hours worked, M_tH_t, the firm can decrease either M_t or HP_t or both.

In the present model the desired distribution of M_tHP_t between M_t and HP_t is assumed to be a function of HS_t. HS_t is the dividing line between standard hours of work and more costly overtime hours: if HP_t is greater than HS_t, then an overtime premium has to be paid on the hours above HS_t. It is thus assumed that the long run equilibrium number of hours paid-for per worker is HS_t. With this in mind, the measure of "excess labor" is taken here to be $\log HS_t - \log H_t$, 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 HS_t, there is considered to be a positive amount of excess labor on hand (i.e., too many workers on hand), and if H_t is greater than HS_t, there is considered to be a negative

---

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

[2]For reasons which 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., $\log HS_t - \log H_t$) will be referred to merely as the difference of the variables.

amount of excess labor on hand (i.e., too few workers on hand).[1] How the amount of excess labor on hand is assumed to affect changes in M_t and HP_t 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 a) no short-run substitution possibilities between workers and machines and b) 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. Because of assumptions a) and b), the number of hours worked per worker, H_t , is also the number of hours worked per machine. The short-run production function is thus:

$$(1) \qquad Y_t = \min \{ \alpha_t M_t H_t , \beta_t K_t H_t \} ,$$

---

[1]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 HS_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, HS_t should be considered to be the desired number of hours paid-for and worked per worker.

where α_t and β_t are coefficients which may be changing through time as a result of technical progress. From the definition of H_t, it is implicitly assumed here (as well as in [3]) that $\alpha_t M_t H_t$ equals $\beta_t K_t H_t$ in (1), so that (1) implies

$$(2) \qquad M_t H_t = \frac{1}{\alpha_t} Y_t \quad .$$

The Measurement of Aggregate Excess Labor

In [3] 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 and not a relationship between seasonally adjusted inputs and a seasonally adjusted output. Unfortunately perhaps, the world of empirical macro economics 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 here 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 here than on those reached in the study of three-digit industries in [3]. The study here in fact should probably be looked upon merely as an attempt to use some of the ideas and conclusions in [3] to develop an equation for predicting changes in aggregate (seasonally adjusted) employment, rather than as an attempt to test various hypotheses about short-run employment demand.

The data which are used here are quarterly data on Y_t, M_t, and HP_t compiled by the Bureau of Labor Statistics for the private nonfarm sector.[1] Data, of course, are not available on H_t. In Figure 1 output per paid-for man hour, Y_t/M_tHP_t, is plotted for the 471-692 period. The dotted lines in the figure are peak to peak interpolation lines of the series.[2] The slopes of the five interpolation lines since 513 are respectively .000001034, .000000882, .000001657, .000001743, and .000001402. From these interpolations, the growth in output per paid-for man hour appears to have been greater for the 601-661 period than for the periods before and after.

The assumption is now made that at each of the interpolation peaks in Figure 1 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 (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 for each quarter of the sample period,[3] which from (2) and from

---

[1]The data on Y_t and M_tHP_t are currently published (in index number form) in the Monthly Labor Review, Table 32. The data used here are not in index number form and were obtained directly from the BLS. The data on Y_t are in units of billions of 1958 dollars, on M_t in units of thousands of workers, and on HP_t in units of hours per week per worker.

[2]The choices for the peaks were, of course, somewhat arbitrary, although the results below were not very sensitive to slightly different choices.

[3]The 661-684 line was extrapolated to get the 691 and 692 values for α.

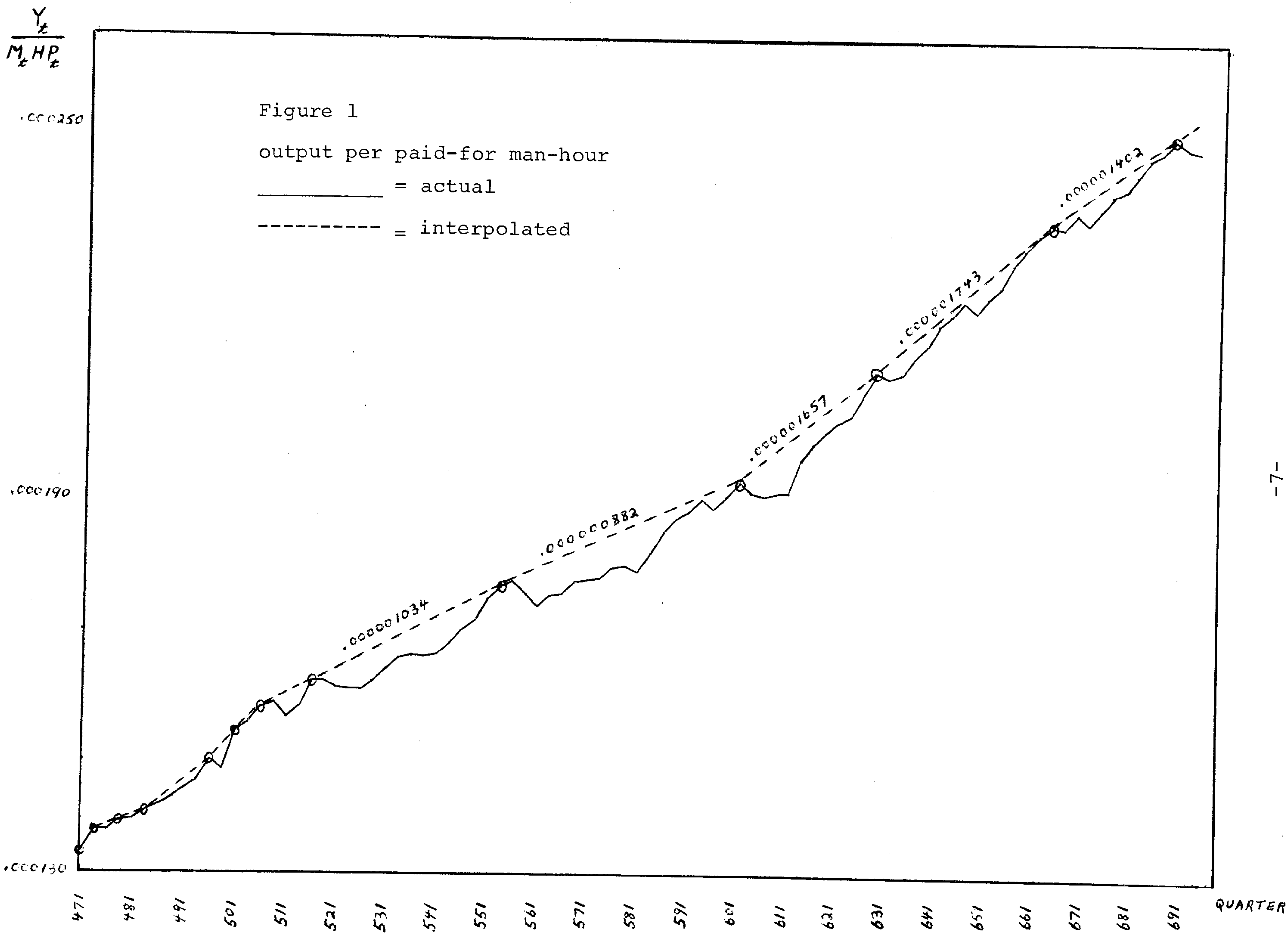

Figure 1

output per paid-for man-hour

the data on Y_t allows an estimate of man-hour requirements, M_tH_t, to be made for each quarter. For any quarter, M_tH_t is the estimated number of man hours required to produce Y_t. Finally, if M_tH_t is divided by HS_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:

$$(3) \qquad M_t^d = M_tH_t/HS_t \quad .$$

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_tH_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 can be 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, as was done in [3], pp. 49-50, that $\log M_t - \log M_t^d$ equals $\log HS_t - \log H_t$, the latter being the measure of excess labor on hand defined at the beginning of this section.

In equation (5) below HS_t is assumed to be a slowly trending variable, the parameters of which are estimated in the regression equation (4)'. Therefore, from the estimate of equation (4)' below (see equation (7)), an estimate of the HS_t series can be derived. The already constructed M_tH_t series can then be divided by this HS_t series to yield from equation (3) a series on M_t^d. These calculations were made, and in Table 1 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 561-692 period.

Using $M_t - M_t^d$ as the measure of excess labor, there were, according to Table 1, 790,000 too many workers employed in 692 for the amount of output produced. This compares with a range of 1,286,000 too few workers in 661 to 1,996,000 too many workers in 581. The numbers on $M_t - M_t^d$ in Table 1 should give the reader a good idea as to what the excess labor situation was like in any one quarter.

The Short-Run Demand for Workers

In [3] 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. Given the nature of the data here, there is little hope of picking up the influence of output expectations more than one quarter ahead, and the equation determining employment demand is taken here to be:

$$(4) \quad \log M_t - \log M_{t-1} = \alpha_1(\log M_{t-1} - \log M_{t-1}^d) + \gamma_0(\log Y_t^e - \log Y_{t-1}).$$

Equation (4) 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 $t-1$ and of the expected change in

TABLE 1

Estimated Values for M_t^d

Quarter	M_t	M_t^d	$M_t-M_t^d$
561	54821	54467	354
562	55095	54741	354
563	54876	54283	593
564	55150	54667	483
571	55314	54937	377
572	55533	54706	826
573	55423	54809	614
574	54876	53593	1283
581	53782	51786	1996
582	53180	51900	1280
583	53563	53153	410
584	54001	54341	-341
591	54548	55036	-489
592	55423	56423	-1000
593	55423	55557	-134
594	55587	55921	-334
601	56244	57024	-780
602	56408	56258	150
603	56189	55452	737
604	55861	54557	1304
611	55751	53983	1768
612	55642	54845	797
613	55916	55542	373
614	56517	56280	238
621	56955	56758	197
622	57338	57230	108
623	57393	57646	-254
624	57174	57863	-689
631	57447	57569	-122
632	57776	57617	159
633	58159	58283	-124
634	58268	58704	-436
641	58706	59353	-648
642	59198	59715	-517
643	59472	60118	-646
644	59909	59919	-9
651	60457	60863	-406
652	61004	61341	-338
653	61606	62263	-658
654	62317	63352	-1035
661	62919	64204	-1286
662	63520	64523	-1002
663	64177	64662	-485
664	64451	65020	-569
671	64724	64331	393
672	64779	64299	480
673	64943	64739	204
674	65381	65005	376
681	65818	65608	210
682	66365	66629	-264
683	66639	66925	-286
684	67022	67260	-238
691	67733	67300	433
692	68171	67381	789

Figures are in thousands of workers.

output for quarter t . α_1 is expected to be negative and γ_o to be positive.

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 is assumed that $Y_t^e = Y_t$ in (4).[1] In other words, output expectations are assumed to be perfect for the current quarter. In [3] the past change in output variable, $\log Y_{t-1} - \log Y_{t-2}$, had a significant coefficient estimate for many of the industries tested, but it is likely that this was due merely to the time period nature of the monthly data used.[2] Otherwise there was little evidence in [3] that past output change variables were significant in an equation like (4), and the $\log M_{t-1} - \log M_{t-1}^d$ term appeared to depict adequately the reaction of firms to the amount of excess labor on hand.

One more assumption is necessary before equation (4) can be estimated. It is assumed, as mentioned above, that the standard number of hours of work per worker, HS , is either a constant or a slowly trending variable, and specifically that

(5) $$HS_{t-1} = \overline{H} e^{\mu t} ,$$

---

[1]As discussed in [3], pp. 5-6, for monthly data the exact time period to which a variable refers is of important consideration. For quarterly data, however, knowledge of the exact period is of less importance, and this issue has largely been ignored in the discussion here.

[2]See the discussion on p. 63 of [3].

where $\bar{H}$ and μ are constants. From this assumption and the others above, equation (4) can be written:

$$(4)' \quad \log M_t - \log M_{t-1} = \alpha_1 \log \bar{H} + \alpha_1 \mu t + \alpha_1(\log M_{t-1} - \log M_{t-1}H_{t-1}) + \gamma_0(\log Y_t - \log Y_{t-1}).$$

Data on $M_{t-1}H_{t-1}$ in (4)' were constructed in the manner described above.

There are perhaps two main differences between equation (4) 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 (4) does not directly include a capital stock variable. It is instead assumed here that there are no short-run substitution possibilities between workers and machines and that the long-run effects of the growth of technical progress (as embodied in, say, new capital stock) on employment are reflected in the movement through time of α_t in (2). If α_t is increasing through time, then, other things being equal, M_t^d in (3) will be falling, since man-hour requirements, M_tH_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 Short-Run Demand for Hours Paid-For Per Worker[1]

In many macro economic models there is a need to determine the number of hours paid-for per worker so that total man-hours paid-for can be determined. The latter is then used in the determination of the total wage bill. One of the findings of [3] was that many of the same factors which determine the change in the number of workers employed also determine the change in the number of hours paid-for per worker. It was also found that the variable, $\log HP_{t-1} - \log HS_{t-1}$, is a significant determinant of the change in hours paid-for per worker. Unlike the number of workers employed, which can move steadily upward or downward over time, the number of hours paid-for per worker fluctuates around a relatively constant level of hours. If, for example, the number of hours paid-for per worker is greater than this level, this should, other things being equal, bring forces into play causing it to decline back to this level. This is the reason for the inclusion of the $\log HP_{t-1} - \log HS_{t-1}$ variable. The final equation determining the change in the number of hours paid-for per worker is thus taken to include the $\log HP_{t-1} - \log HS_{t-1}$ variable, the amount of excess labor on hand, and the current change in output:

$$(6) \quad \log HP_t - \log HP_{t-1} = \alpha_1'(\log M_{t-1} - \log M_{t-1}^d) + \alpha_2'(\log HP_{t-1} - \log HS_{t-1}) + \gamma_0'(\log Y_t^e - \log Y_{t-1}) \quad ,$$

[1]See [3], pp. 139-144, for a more detailed development of the theoretical model of the short-run demand for hours paid-for per worker.

or on the above assumptions:

$$(6)' \quad \log HP_t - \log HP_{t-1} + (\alpha_1' - \alpha_2')\log \overline{H} + (\alpha_1' - \alpha_2')\mu t + \alpha_1'(\log M_{t-1} - \log M_{t-1}H_{t-1}) + \alpha_2' \log HP_{t-1} + \gamma_0'(\log Y_t - \log Y_{t-1}) \quad .$$

III. ESTIMATES OF THE EMPLOYMENT DEMAND MODEL

The results of estimating equation (4)' for the 561-692 period are:[1]

$$(7) \quad \log M_t - \log M_{t-1} = \underset{(7.43)}{-1.022} + \underset{(5.56)}{.000172t} - \underset{(7.40)}{.277}(\log M_{t-1} - \log M_{t-1}H_{t-1}) + \underset{(9.47)}{.356}(\log Y_t - \log Y_{t-1}), \; R^2 = .773, \; SE = .00316, \; DW = 1.41, \; 54 \text{ obs.}$$

All of the estimates in (7) are highly significant and of the expected sign. The estimate of the coefficient of the excess labor variable is -.277, which implies that, other things being equal, about 28 percent of the amount of excess labor is removed each quarter. The equation was also estimated for the longer 473-692 period, with the following results:

$$\widehat{\alpha_1 \log \overline{H}} = \underset{(7.28)}{-.980}, \quad \widehat{\alpha_1 \mu} = \underset{(6.80)}{.000231}, \quad \hat{\alpha}_1 = \underset{(7.27)}{-.265}, \quad \hat{\gamma}_0 = \underset{(11.67)}{.415},$$

$R^2 = .702$, SE = .00459, DW = 1.40, 88 obs. The estimates for the

---

[1]t-statistics (in absolute values) are in parentheses. In what follows a coefficient estimate will be said to be "significant" if its t-statistic is greater than two in absolute value and a variable will be said to be "significant" if its coefficient estimate is significant.

two periods are not vastly different, but a Chow test of the hypothesis that the coefficients of the equation are the same for the 473-554 period as for the 561-692 period yielded negative results ($F_{4,80} = 4.38$). From this test the coefficient of $\log Y_t - \log Y_{t-1}$ appeared to be larger for the earlier period than for the later one. The 561-692 period was thus chosen as the basic period of estimation here, although this decision was perhaps somewhat arbitrary.

Equation (7) was also estimated under the assumption of first order serial correlation of the error terms, with the following results:[1]

$$(8) \quad \log M_t - \log M_{t-1} = \underset{(6.55)}{-1.085} + \underset{(4.66)}{.000185t} - \underset{(6.52)}{.294}(\log M_{t-1} - \log M_{t-1}H_{t-1})$$

$$+ \underset{(8.95)}{.348}(\log Y_t - \log Y_{t-1}), \; \hat{\rho} = .275, \; SE = .00305, \; 54 \text{ obs.}$$

The coefficient estimates in (8) are little changed from those in (7) and the estimate of the serial correlation coefficient ρ is fairly small at .275. Serial correlation of the residuals thus does not appear to be a serious problem in (7).

A few other equations similar to (7) were also estimated. $\log Y_{t-1} - \log Y_{t-2}$ was added to the equation, and as expected for the quarterly data here, 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 (7) (under

[1]The equation was estimated by the Cochrane-Orcutt iterative technique with a tolerance level of .005 between successive estimates of ρ.

the hypothesis of perfect expectations), and it likewise was not significant.[1] 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 (7) 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 (7) merely because it is of the nature of a lagged dependent variable.[2] The results were quite poor and $\log M_{t-1}$ 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 (7).[3] Equation (7) was thus chosen as the basic equation determining the change in the number of workers employed.

Turning to the hours equation, the results of estimating equation (6)' for the 561-692 period are:

$$(9) \quad \log HP_t - \log HP_{t-1} = \underset{(3.58)}{.795} - \underset{(3.70)}{.000210t} - \underset{(2.55)}{.102}(\log M_{t-1} - \log M_{t-1}H_{t-1}) - \underset{(3.97)}{.315}\log HP_t + \underset{(4.26)}{.125}(\log Y_t - \log Y_{t-1}),$$

$R^2 = .398$, SE = .00246, DW = 2.07, 54 obs.

---

[1]The equation included only 53 observations, since the 692 observation had to be dropped to allow for the last observation for Y_{t+1}.

[2]See the more complete discussion of this in [3], pp. 72-76.

[3]$R^2 = .724$ vs. .773 in (7). Also, as argued in [3], 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.

All of the coefficient estimates are significant in (9), and in particular the coefficient of the excess labor variable is significant. The change in aggregate hours paid-for per worker does appear to be determined in part by the amount of aggregate excess labor on hand. A similar conclusion was reached in [3] for much more disaggregate data.

Equation (9) was also estimated under the assumption of first order serial correlation of the error terms. There was little change in the coefficient estimates and the estimate of the serial correlation coefficient was only -.100. Serial correlation thus does not appear to be a problem in (9). In [3] it was argued[1] that the unemployment rate (as a measure of the condition of the labor market) is likely to have a negative effect on $\log HP_t - \log HP_{t-1}$, and the results for the three-digit industries indicated that this is the case. For the aggregate data here, however, when $\log UR_t$ (where UR_t is the civilian unemployment rate) was added to (9), it was not significant. Equation (9) was thus taken as the basic equation determining the change in hours paid-for per worker.

Adding equations (7) and (9) yields an equation determining the change in total man-hours paid-for:

$$(10)\quad \log M_t HP_t - \log M_{t-1} HP_{t-1} = -.227 - .000038t - .379(\log M_{t-1} - \log M_{t-1} H_{t-1}) - .315 \log HP_{t-1} + .481(\log Y_t - \log Y_{t-1}) \;.$$

---

[1]See [3], pp. 142-143.

Notice that the coefficient of $\log Y_t - \log Y_{t-1}$ is much less than one, which implies that, say, a one percent increase in output leads to a less than one percent increase in man-hours paid-for in the current period.[1]

IV. LINK BETWEEN EMPLOYMENT PREDICTIONS AND PREDICTIONS OF THE UNEMPLOYMENT RATE

It is not a trivial matter to go from predictions of M_t to predictions of the unemployment rate. First of all 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 unemployment rate; and thirdly M_t provides no direct projections of the labor force. In this section a somewhat crude attempt is made to link the predictions of M_t in Section III to predictions of the unemployment rate. From the nature of the work in this section less reliance should be put on the results and conclusions reached here than on those reached in Section III.

Let MA_t denote the number of agricultural workers employed during quarter t, MG_t the number of government workers employed, and E_t the total number of civilian workers employed according to the household survey. Then D_t is defined to be

---

[1] Notice also that the coefficient of $\log M_{t-1} - \log M_{t-1}H_{t-1}$ is not equal to the coefficient of $\log HP_{t-1}$ in (10). See [3], p. 161, for a discussion of the implications of this for the specification of the total man-hours paid-for equation.

(11) $$D_t = M_t + MA_t + MG_t - E_t \quad .$$

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

D_t appears to respond to labor market conditions, and the following equation was estimated for the 482-692 period[1] under the assumption of first order serial correlation of the error terms:

(12) $$D_t = \underset{(6.65)}{-14285} - \underset{(4.83)}{58.03t} + \underset{(8.10)}{.385}\, M_t \ , \quad \hat{\rho} = .799, \ SE = 222.6, \ 85 \text{ obs.}$$

What (12) says is that, other things being equal, a change in M_t of, say, 1000 leads to a change in E_t of only 615. The difference of 385 is taken up either by moonlighters or by other discrepancies between the establishment and household surveys.[2] Notice that the estimate of the serial correlation coefficient is fairly high in equation (12), which may indicate that the specification of the equation is too simple.

---

[1]For most of the work in this section the estimation period was taken to be 482-692, rather than the shorter 561-692 period used in the previous section.

[2]Black and Russell [1] have estimated an equation similar to (12), using, however, the unemployment rate in place of M_t. The Black and Russell equation was also estimated for the work here, but it led to poorer results than those in (12).

Once M_t is determined, D_t can be determined, and then taking MA_t and MG_t as exogenous in (11), E_t can be determined. This leaves only the labor force to be determined in order to calculate the unemployment rate. There are many special factors which are likely to affect labor force participation rates--some of which have been ably described by Mincer [4]--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 results here are thus not meant to provide answers to current questions about the determinants of participation rates; it is merely hoped that the estimated equations will provide a reasonably good link from predictions of aggregate employment to predictions of 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:

$$(13) \quad \frac{LF_{1t}}{P_{1t}} = \underset{(721.59)}{.9804} \; \underset{(8.38)}{-.000174t}, \quad \hat{\rho} = .169, \; SE = .00197, \; 54 \text{ obs.}$$

LF_{1t} denotes the primary (males 25-54) labor force (including armed forces) during quarter t, and P_{1t} denotes the non-institutional population (including armed forces) of males 25-54 during quarter t. Equation (13) was estimated for the 561-692 period under the assumption of first order serial correlation of the error terms. LF_{1t}/P_{1t} rose from 1948 to the mid-fifties and then slowly declined from the mid-fifties to the present, and this is the reason the shorter 561-692 period was used to estimate equation (13).

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 which 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 to 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:

$$(14) \quad \frac{LF_{2t}}{P_{2t}} = \underset{(5.10)}{.2819} + \underset{(4.27)}{.000401t} + \underset{(2.95)}{.2838}\,\frac{E_t + AF_t}{P_{1t} + P_{2t}}, \quad \hat{\rho} = .883,$$

SE = .00254, 85 obs.

LF_{2t} denotes the secondary labor force (including armed forces) during quarter t, P_{2t} the non-institutional population (including armed forces) of everyone over 16 except males 25-54, and

AF_t the level of the total armed forces. Equation (14) was estimated for the 482-692 period.

There is one obvious statistical problem in estimating an equation like (14), 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 (14) will thus be biased upward unless account is taken of the errors of measurement problem. Equation (14) was thus estimated by an instrumental variable or two stage least squares technique, with account also being taken of the serial correlation of the error terms. The technique which was used is described in [2]. Basically, the technique is a combination of the Cochrane-Orcutt iterative technique and the two stage least squares technique.

The instruments which were used for $E_t+AF_t/P_{1t}+P_{2t}$ in (14) are the constant, t, LF_{2t-1}/P_{2t-1}, $E_{t-1}+AF_{t-1}/P_{1t-1}+P_{2t-1}$, $AF_t/P_{1t}+P_{2t}$, $M_t/P_{1t}+P_{2t}$, $M_{t-1}/P_{1t-1}+P_{2t-1}$, $MG_t/P_{1t}+P_{2t}$, and $MA_{t-1}+MG_{t-1}/P_{1t-1}+P_{2t-1}$. As discussed in [2], the first four instruments are necessary in order to insure consistent estimates. The other instruments are based on equations (11) and (12). Write equation (12) as

$$(12)' \qquad D_t = \psi_o + \psi_1 t + \psi_2 M_t + \mu_t ,$$

where $\mu_t = \rho_o \mu_{t-1} + \varepsilon_t$. (The error term ε_t is assumed to have mean zero and constant variance and to be uncorrelated with M_t and with its own past values.) Combining (11) and (12)' and solving for E_t yields:

$$(15) \quad E_t = -\psi_0(1+\rho_0) + \rho_0\psi_1 - \psi_1(1+\rho_0)t + (1-\psi_2)M_t + MA_t + MG_t - \rho_0 E_{t-1} + \rho_0(1-\psi_2)M_{t-1} + \rho_0 MA_{t-1} + \rho_0 MG_{t-1} - \varepsilon_t .$$

Since M_t, M_{t-1}, MG_t, and MG_{t-1} in equation (15) 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 (15) will not be correlated with the measurement error in E_t and thus can be used as an instrument.

When an equation like (14) 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 .4982, which seems to indicate that, unless corrected, the measurement error bias is quite large in equations like (14). The same conclusion was also reached when different periods of estimation were used, although as the period of estimation was made smaller by moving the initial observation closer to the present, the coefficient estimates under both techniques increased in size.

Equation (14) is similar to equations estimated by Tella [4], [5], 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 (14) appeared to give the best results. In general, the results achieved here are only moderate, as is perhaps best indicated by the large amount of serial correlation in (14).

Having determined E_t and taking P_{1t}, P_{2t}, and AF_t as exogenous, LF_{1t} and LF_{2t} can be determined from equations (13) and (14). By definition the civilian labor force (denoted as CLF_t) is equal to $LF_{1t} + LF_{2t} - AF_t$, and so having determined CLF_t , the unemployment rate can be determined as $UR_t = 1 - E_t/CLF_t$.

V. SIMULATION OF THE OVERALL MODEL

In order to gauge the accuracy of the overall model, it was simulated, both statically and dynamically, for the 561-692 period. Equation (7) was first used to determine M (Y and the interpolation lines in Figure 1 taken to be exogenous); equation (12) was then used to determine D ; equation (11) was next used to determine E (MA and MG taken to be exogenous); and finally equations (13) and (14) were used to determine LF_1 and LF_2 (P_1, P_2 , and AF taken to be exogenous). The unemployment rate was then computed as $1 - E/(LF_1 + LF_2 - AF)$.

For the static simulations the actual values of the lagged endogenous variables in the model were used, whereas for the dynamic simulations the generated values of these variables were used (aside from the initial values for 554). Because of the serial correlation of the residuals, lagged endogenous variables appear in equations (12), (13), and (14), as well as in the employment demand equation (7). For each static simulation the mean absolute error in terms of levels was computed for each of the endogenous variables and for each dynamic simulation the mean absolute error in terms of both levels and changes was computed.[1]

In Table 2 the predicted values of M_t are presented for both the static and dynamic simulations of equation (7). For the static simulation the mean absolute error is 136.2 thousand workers, and for the dynamic simulation the mean absolute error is 233.9 thousand workers in terms of levels and 142.7 thousand workers in terms of changes. Overall, the results in Table 2 look fairly good. The decreases in employment

---

[1]Let $\hat{X}_t$ denote the predicted value of X_t. The mean absolute error in terms of levels (MAE) and the mean absolute error in terms of changes (MAEΔ) are defined to be:

$$MAE = \frac{\sum_{t=1}^{T} |\hat{X}_t - X_t|}{T}, \quad MAE\Delta = \frac{\sum_{t=1}^{T} |(\hat{X}_t - \hat{X}_{t-1}) - (X_t - X_{t-1})|}{T},$$

where T is the number of observations. MAEΔ has the advantage of eliminating the compounding of the level errors.

TABLE 2

Actual and Simulated Values of M_t

Quarter	M_t	Static Simulation $\hat{M}_t$	Static Simulation $\hat{M}_t - M_t$	Dynamic Simulation $\hat{M}_t$	Dynamic Simulation $\hat{M}_t - M_t$
561	54821	54437	-384	54437	-384
562	55095	54976	-119	54697	-397
563	54876	54987	111	54700	-176
564	55150	55003	-146	54876	-274
571	55314	55267	-47	55068	-245
572	55533	55281	-252	55104	-429
573	55423	55494	71	55184	-239
574	54876	54965	89	54794	-82
581	53782	54005	224	53947	165
582	53180	53410	230	53529	349
583	53563	53420	-143	53673	110
584	54001	54020	19	54100	99
591	54548	54489	-58	54562	14
592	55423	55320	-103	55331	-92
593	55423	55544	121	55477	54
594	55587	55741	153	55780	193
601	56244	56220	-24	56361	117
602	56408	56420	12	56505	97
603	56189	56308	119	56378	189
604	55861	55888	27	56024	163
611	55751	55512	-239	55629	-122
612	55642	55792	150	55703	62
613	55916	55893	-23	55938	22
614	56517	56298	-219	56314	-203
621	56955	56846	-109	56698	-257
622	57338	57293	-45	57106	-232
623	57393	57681	288	57512	120
624	57174	57764	590	57851	677
631	57447	57490	43	57981	534
632	57776	57728	-48	58115	339
633	58159	58198	40	58445	286
634	58268	58572	304	58781	513
641	58706	58848	142	59222	516
642	59198	59242	44	59618	420
643	59472	59714	242	60020	548
644	59909	59809	-100	60207	298
651	60457	60477	21	60694	238
652	61004	60970	-34	61143	139
653	61606	61655	50	61757	151
654	62317	62404	87	62515	198
661	62919	63136	218	63281	363
662	63520	63588	67	63852	332
663	64177	64050	-127	64292	115
664	64451	64643	192	64726	276
671	64724	64568	-156	64767	43
672	64779	64808	29	64839	60
673	64943	65007	64	65050	108
674	65381	65185	-196	65263	-118
681	65818	65696	-122	65611	-208
682	66365	66328	-37	66177	-188
683	66639	66750	111	66613	-26
684	67022	67043	21	67024	2
691	67733	67309	-425	67310	-423
692	68171	67850	-321	67543	-628

Figures are in thousands of workers.

MAE = 136.2 (Static Simulation)

MAE = 233.9 (Dynamic Simulation)
MAEΔ = 142.7

in the late 57 - early 58 period and the late 60 - early 61 period were caught reasonably well, as were the large increases in the 64-66 period. The model underestimated, however, the increase in employment in 691 and 692. In 691 and 692 the rate of growth of real output declined, but employment did not appear to respond. Notice from Table 1 that this caused the amount of excess labor on hand to be fairly large by 692.

The other key endogenous variable in the model beside M_t is the secondary labor force, and in Table 3 the predicted values of this variable are presented for the static and dynamic simulations. The results in Table 3 are less encouraging. The static simulation results are reasonable, but for the dynamic simulation results the mean absolute error in terms of levels is quite large at 416.3 thousand workers. In particular, the model consistently overpredicted the size of the labor force in the 62 - 65 period and underpredicted the size in the 67 - 69 period. As mentioned in Section IV, other labor force participation equations were tried beside equation (14), but none proved to be any more successful. In short, the aggregate type of participation equations estimated in this study do not appear to be capable of accounting very well for the slow growth of the secondary labor force in the early 60's and the much more rapid growth in the late 60's.

With respect to predictions of the unemployment rate, there is some error cancellation in the model. Positive errors in predicting M_t , for example, will lead, other things being equal, to positive errors in predicting D_t (from equation (12)), 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 the

TABLE 3

Actual and Simulated Values of LF_{2t}

		Static Simulation		Dynamic Simulation	
Quarter	LF_{2t}	$\hat{LF}_{2t}$	$\hat{LF}_{2t}-LF_{2t}$	$\hat{LF}_{2t}$	$\hat{LF}_{2t}-LF_{2t}$
561	37788	37866	78	37866	78
562	38129	37921	-208	37997	-132
563	38233	38193	-40	38081	-152
564	38114	38295	181	38165	52
571	38076	38257	180	38308	232
572	38100	38231	130	38440	340
573	38332	38294	-38	38598	266
574	38427	38404	-23	38640	213
581	38376	38423	47	38609	233
582	38723	38475	-247	38671	-52
583	38799	38892	93	38834	35
584	38712	39048	336	39070	357
591	38829	38985	156	39295	466
592	39121	39203	82	39611	490
593	39325	39306	-19	39737	412
594	39578	39577	-1	39941	363
601	39452	39916	463	40231	779
602	40328	39744	-584	40427	99
603	40479	40533	55	40624	145
604	40743	40629	-114	40754	11
611	41140	40865	-275	40871	-269
612	41321	41291	-30	41054	-267
613	40910	41451	541	41215	305
614	40990	41198	207	41462	472
621	41218	41311	93	41727	509
622	41274	41524	250	41973	699
623	41521	41590	69	42207	686
624	41521	41878	357	42482	961
631	42000	41903	-97	42742	742
632	42415	42342	-72	42988	574
633	42655	42758	103	43263	608
634	42856	43017	162	43553	698
641	43090	43199	109	43810	720
642	43735	43467	-269	44100	365
643	43665	44078	413	44403	738
644	43860	44000	139	44648	788
651	44256	44220	-36	44916	660
652	44731	44672	-59	45260	529
653	45059	45078	19	45549	489
654	45415	45460	45	45896	481
661	45746	45821	75	46248	502
662	46189	46152	-37	46593	404
663	46793	46556	-237	46904	111
664	47362	47151	-211	47244	-118
671	47631	47594	-38	47482	-149
672	47782	47847	65	47715	-67
673	48548	48074	-474	48010	-537
674	48964	48770	-194	48296	-668
681	48999	49168	169	48582	-417
682	49431	49224	-208	48859	-572
683	49550	49635	85	49132	-418
684	49693	49780	87	49410	-283
691	50582	49949	-633	49699	-882
692	50856	50731	-125	49967	-889

Figures are in thousands of workers.

Static Simulation: MAE = 167.7

Dynamic Simulation: MAE = 416.3, MAEΔ = 166.4

secondary labor force (from equation (14)), which will in turn lead to smaller errors in predicting the unemployment rate. In Table 4 the predicted values of the unemployment rate are presented for various simulations. Two static simulations of UR_t are presented in the table, one in which the labor force was taken to be exogenous and the other in which the labor force was treated as endogenous. Two dynamic simulations of UR_t are also presented, again corresponding to the two assumptions about the labor force.

The static simulations of UR_t are quite good, with mean absolute errors of only .00190 and .00199. The high unemployment rates in 58 and 61 were caught fairly well, as were the low unemployment rates in the late 60's. The dynamic simulation results are not as good, and in large part the errors can be traced to errors made in predicting the labor force. When the labor force is treated as exogenous, the mean absolute error in terms of levels is a respectable .00296 and the overall results look fairly good, but with the labor force treated as endogenous the mean absolute error rises to .00438.

For this latter simulation, the model consistently overpredicted the unemployment rate in the 62 - 65 period and underpredicted it in the 663 - 692 period. In 691 there was a huge increase in both the level of employment and the size of the labor force, which the model only in part accounted for. Therefore, for the two simulations in which the labor force was taken to be exogenous, there was no error cancellation, and the predictions of the unemployment rate for 691 were much too

TABLE 4

Actual and Simulated Values of UR_t

Quarter	UR_t	Static Simulation LF exog. $\hat{UR}_t$	Static Simulation LF endog. $\hat{UR}_t$	Dynamic Simulation LF exog. $\hat{UR}_t$	Dynamic Simulation LF endog. $\hat{UR}_t$
561	.0404	.0438	.0434	.0438	.0434
562	.0420	.0408	.0380	.0432	.0413
563	.0414	.0433	.0435	.0439	.0425
564	.0411	.0418	.0450	.0437	.0452
571	.0396	.0394	.0412	.0412	.0439
572	.0407	.0387	.0403	.0398	.0442
573	.0422	.0411	.0411	.0401	.0443
574	.0494	.0487	.0495	.0468	.0510
581	.0629	.0572	.0600	.0551	.0607
582	.0737	.0723	.0682	.0663	.0653
583	.0732	.0719	.0713	.0662	.0647
584	.0636	.0644	.0685	.0589	.0629
591	.0581	.0577	.0605	.0540	.0609
592	.0512	.0502	.0513	.0470	.0537
593	.0530	.0555	.0552	.0520	.0577
594	.0561	.0525	.0523	.0517	.0566
601	.0517	.0505	.0573	.0472	.0583
602	.0524	.0587	.0506	.0552	.0565
603	.0556	.0519	.0523	.0542	.0558
604	.0627	.0623	.0598	.0613	.0605
611	.0678	.0691	.0666	.0681	.0656
612	.0698	.0700	.0689	.0701	.0660
613	.0676	.0633	.0704	.0636	.0675
614	.0619	.0616	.0645	.0584	.0648
621	.0562	.0554	.0577	.0525	.0604
622	.0549	.0545	.0574	.0513	.0604
623	.0555	.0551	.0555	.0521	.0606
624	.0552	.0520	.0567	.0494	.0619
631	.0578	.0560	.0562	.0518	.0630
632	.0569	.0601	.0601	.0557	.0643
633	.0551	.0553	.0576	.0545	.0636
634	.0559	.0535	.0551	.0532	.0619
641	.0546	.0553	.0560	.0534	.0620
642	.0523	.0569	.0529	.0563	.0603
643	.0502	.0456	.0511	.0491	.0586
644	.0497	.0502	.0521	.0496	.0599
651	.0487	.0511	.0503	.0512	.0593
652	.0467	.0465	.0451	.0487	.0547
653	.0438	.0448	.0450	.0465	.0526
654	.0412	.0401	.0408	.0423	.0486
661	.0383	.0366	.0369	.0377	.0434
662	.0383	.0340	.0335	.0337	.0387
663	.0379	.0378	.0350	.0343	.0360
664	.0369	.0373	.0341	.0344	.0325
671	.0373	.0393	.0367	.0375	.0334
672	.0386	.0359	.0368	.0361	.0349
673	.0390	.0433	.0379	.0414	.0351
674	.0392	.0409	.0387	.0428	.0349
681	.0366	.0363	.0374	.0392	.0331
682	.0360	.0361	.0342	.0380	.0315
683	.0360	.0350	.0365	.0364	.0320
684	.0340	.0346	.0364	.0349	.0324
691	.0332	.0424	.0336	.0432	.0314
692	.0349	.0335	.0339	.0412	.0322
		MAE = .00190	MAE = .00199	MAE = .00296	MAE = .00438
				MAEΔ = .00204	MAEΔ = .00195

high. For the dynamic simulation this error also fed into the prediction for 692.

Overall, the unemployment rate results look encouraging. For the dynamic simulation with the labor force treated as endogenous the level errors tend to compound, but even here the mean absolute error in terms of changes is only .00195. The static simulation results are quite good, and there seems to be little problem in predicting the unemployment rate one quarter ahead. The model was also simulated five quarters at a time, with the base quarter being increased by one after each five quarter simulation. The mean absolute errors were then computed for all of the second-quarter predictions, all of the third-quarter predictions, all of the fourth-quarter predictions, and all of the fifth-quarter predictions. The mean absolute errors (in terms of levels) for the second- and third-quarter predictions were close to those for the first-quarter (static simulation) predictions, and only for the fourth- and fifth-quarter predictions were the errors closer to the dynamic simulation errors.

In summary, if the labor force can be predicted adequately, the employment demand model in Section II appears to be capable of predicting the unemployment rate quite well. If, on the other hand, the labor force continues to pose problems for prediction purposes, the model appears to be capable of predicting the unemployment rate fairly well about three quarters in advance, but, except in terms of changes, less so after that. All of these conclusions, of course, are predicated on the assumption that real output and the other exogenous variables of the model can be adequately predicted.

REFERENCES

[1] Black, Stanley W., and R. Robert Russell, "An Alternative Estimate of Potential GNP," The Review of Economics and Statistics, LI (February 1969), 70-76.

[2] Fair, Ray C., "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors," Econometrica, forthcoming.

[3] ________, The Short-Run Demand for Workers and Hours, (Amsterdam: North-Holland Publishing Company, 1969).

[4] Mincer, Jacob, "Labor-Force Participation and Unemployment: A Review of Recent Evidence," in R. A. Gordon and M. S. Gordon, eds., Prosperity and Unemployment, (New York: John Wiley, 1966), 73-112.

[5] Tella, A. J., "The Relation of Labor Force to Employment," Industrial and Labor Relations Review, 17 (April 1964), 454-469.

[6] ________, "Labor Force Sensitivity to Employment by Age, Sex," Industrial Relations, 4 (February 1965), 69-83.

[7] Nerlove, Marc, "Notes on the Production and Derived Demand Relations Included in Macro-Econometric Models," International Economic Review, 8 (June 1967), 223-242.