CHAPTER 4

THE BASIC RESULTS FOR PRODUCTION WORKERS

4.1 Introduction

In this chapter the results are presented of estimating eq. (3.9) under the expectational hypotheses discussed above. These basic results are discussed in detail in this chapter, and then in ch. 5, using eq. (3.9) as a starting point, various hypotheses regarding other possible determinants of short-run fluctuations in the number of workers employed are presented and tested. The data used in this study are considerably more detailed than the data used in previous studies, and these data will be discussed first.

4.2 The data

The basic model of previous studies discussed in ch. 2 and the model developed in this study take the firm as the basic behavioral unit. Data are not available by firm, however, and some amount of aggregating must be done. In many of the previous studies highly aggregate data have been used, such as for all of manufacturing. The use of highly aggregate data is likely to conceal certain relationships which may exist in the disaggregate data, and Hultgren has discovered in his work that the use of large statistical aggregates tends to conceal the disaggregate relationships between fluctuations in output and output per man hour.[1] The basic reason for this is that production cycles in different industries do not coincide with one another and to some extent tend to cancel each other out. Also, one might expect that hiring and firing practices would differ considerably across industries.

The two studies of the United States which have used quarterly two-digit industry data are those of DHRYMES (1967) and KUH (1965b). Unfortunately, much of these data are nearly useless for the study of short-run relationships between output and employment. The quarterly two-digit industry data have been constructed by interpolating annual two-digit industry data using

---

1 HULTGREN (1960, pp. 28–29).

the Federal Reserve Board (FRB) indices of industrial production. About half (by value added) of the FRB indices, however, are obtained by interpolating annual data using the Bureau of Labor Statistics (BLS) man-hour data and some assumption about how output per man hour fluctuates with output in the short run – and this is one of the very things the studies are concerned with estimating. When these data are used, combined with the BLS data on employment, to estimate the relationship between employment and output in the short run, the net result is to estimate the estimating technique used by the FRB to construct the output data in the first place. There are only four two-digit industries in which the data are not based at least in part on man-hour interpolations – 33 Primary metals, 26 Paper and allied products, 21 Tobacco manufacturing, and 29 Petroleum refining and related industries.

Fortunately, there are better United States data available, at a sacrifice, however, of complete coverage of all of United States manufacturing. There are seventeen three-digit industries for which FRB output data and BLS employment data are available monthly from 1947 to the present where

TABLE 4.1

The seventeen industries considered in this study

SIC number	Description
201	Meat products
207	Confectionery and related products
211	Cigarettes
212	Cigars
231	Men's and boys' suits and coats
232	Men's and boys' furnishings
233	Women's, misses', and juniors' outerwear
242	Sawmills and planing mills
271	Newspaper publishing and printing
301	Tires and inner tubes
311	Leather tanning and finishing
314	Footwear, except rubber
324	Cement, hydraulic
331	Blast furnace and basic steel products
332	Iron and steel foundries
336	Non-ferrous foundries
341	Metal cans

the FRB output data are measured independently of BLS employment data. In addition there are about twenty four-digit industries for which these data are available monthly from 1958 to the present.[1] The seventeen three-digit manufacturing industries considered in this study are listed in table 4.1. These industries constitute about eighteen percent of manufacturing by value added.

There are other advantages of using these data in addition to the output estimates being independent of the man-hour data. For the three-digit industries the degree of disaggregation is quite good, and many of the problems with using highly aggregate data should be mitigated. The three-digit industries are much more homogeneous groups than even the two-digit industries. The use of monthly data in a short-run study also seems very desirable, as some of the relationships between short-run fluctuations in employment and output may be covered up in quarterly data.

The BLS production worker data used in this study refer to persons on establishment pay-rolls who receive pay for any part of the pay period which includes the 12th of the month. Persons who are on paid sick leave, on paid holidays and vacations, or who work during part of the pay period and are on strike or unemployed during the rest of the period are counted as employed. These are the data which were used for M_{2wt}. Data for the average number of hours paid-for per production worker during the second week of month t, HP_{2wt}, were also taken from the BLS. These data are compiled from the same survey as the data on the number of production workers employed.

The FRB data on output do not refer to the amount of output produced during the second week of the month, which from the point of view of the model developed in ch. 3 it would be desirable to have, but instead refer to the average daily rate of output for the month. For lack of a better alternative, however, the FRB data were used as the output data in this study. Unless the weekly rate of output fluctuates considerably during the month, the average daily rate of output for the month should be a fairly good approximation to the average daily rate of output for the second week of the month. The observed average daily rate of output for the month will be denoted as Y_{dt} to distinguish it from the unobserved but theoretically preferred Y_{2wt} variable. Some of the consequences of using Y_{dt} in place of Y_{2wt} will be discussed below.

---

[1] There are also three three-digit mining industries for which data are available from 1947 to the present.

The basic period of estimation was taken to be the period 1947–1965, but a number of adjustments in this basic period were made for each industry. For example, in all of the seventeen industries except 201, 271, 324, and 341, a significant percentage of firms shut down for vacations in July (usually the first two weeks), and in industries 207, 211, 212, 231, 232, 233, and 314 a significant number of firms also shut down during the Christmas week in December.[1] In July and December many of these firms find demand at low levels anyway, and they find it to their advantage to shut the entire plant down for a week or two for vacations, rather than to keep the plant open and spread the vacations over a longer period of time. For these shutdown periods production is clearly not exogenous, and thus it was decided to exclude from the periods of estimation the months in which shutdowns occurred. This means, for example, that for industries which shut down in July and December the values of $\log M_{2wt} - \log M_{2wt-1}$ for June to July, July to August, November to December, and December to January were excluded.

Since past and expected future output changes are assumed in the model developed above to have an effect on employment decisions, excluding the four July and December observations when shutdowns occur does not exclude the July and December output figures from entering the estimated equation. With respect to this problem, the use of the FRB data on the average daily rate of output for the month is probably more desirable than the use, if they were available, of data on the amount of output produced during the second week of the month. If firms shut down during the first two weeks of July, for example, little if any output will be produced during the second week, but the total effect on the average daily rate of output for the month will be less. Regarding the past and expected future output streams, for the months in which shutdowns occur firms are more likely to look at the average daily rate of output for the month than the rate during the second week. To the extent, however, that without the shutdown in, say, July, the average daily rate for July would have been larger and the June and August rates smaller, the $\log Y^e_{d\text{July}} - \log Y^e_{d\text{June}}$ and $\log Y^e_{d\text{Aug.}} - \log Y^e_{d\text{July}}$ variables are probably inadequate measures of the time stream of expected future output changes. This (hopefully slight) misspecification

[1] This information was gathered mainly from industry and union officials.

should be kept in mind when interpreting the estimates of the coefficients of the expected future output variables.[1]

In the development of the model in § 3.7 it was argued that the variables, $\log Y_{2wt-i} - \log Y_{2wt-i-1}$ ($i = 1, 2, \ldots, m$), might be significant in the determination of $\log M_{2wt} - \log M_{2wt-1}$. It was argued that these past output change variables might be picking up part of the reaction of the firm to the amount of excess labor on hand. Because of the fact that the output variable used in the empirical work is not the amount of output produced during the second week of the month, but is rather the average daily rate of output for the month, there is an additional reason why $\log Y_{dt-1} - \log Y_{dt-2}$ may be a significant determinant of $\log M_{2wt} - \log M_{2wt-1}$ in addition to it possibly picking up part of the reaction to the amount of excess labor on hand. Remember that $\log M_{2wt} - \log M_{2wt-1}$ is the change in the number of production workers employed from the second week of month $t - 1$ to the second week of month t. To the extent that output is, say, increasing throughout month $t - 1$ and to the extent that the number of production workers employed responds to this increase during the last half of month $t - 1$, $\log M_{2wt} - \log M_{2wt-1}$ will be influenced by the increase in output during the last two weeks of month $t - 1$. It will therefore be influenced by $\log Y_{dt-1} - \log Y_{dt-2}$, since an increase in output in the last two weeks of month $t - 1$ raises the average daily rate for the whole month, Y_{dt-1}. There is no way these two possible effects can be separated, and so the estimates of the coefficient of $\log Y_{dt-1} - \log Y_{dt-2}$ must be interpreted with caution.

In eight of the seventeen industries there were significant strikes (involving 10,000 workers or more) during the nineteen-year period of estimation. In table 4.2 these strikes are listed by industry, and the date of each strike and the number of workers involved are given. In the actual regressions these observations were omitted, as well as the observations for the two or three months before and after the strike. Some of the estimated equations had lags in output of up to 13 months, however, and either all of these observations had to be omitted or an output variable had to be constructed to use

[1] For the Newspaper industry, 271, the December observations were omitted in the empirical work, since the average daily rate of output for this month was much lower than the rate during the second week, due to the heavy advertising before Christmas and the much lighter advertising after Christmas. There is still a problem here, of course, to the extent that the use of the December output observations is resulting in a bad approximation to the time stream of expected future output changes.

TABLE 4.2

Strikes involving 10,000 or more workers

Industry	Approximate period of strike		No. of workers involved
201	March 16, 1948	June 5, 1948	83,000
	September 4, 1959	October 24, 1959	18,000
233	February 17, 1948	February 19, 1948	10,000
242	April 29, 1952	May 31, 1952	45,000
	June 21, 1954	September 13, 1954	77,000
	June 5, 1963	August 18, 1963	29,000
271	December 8, 1962	March 31, 1963	20,000
	September 16, 1965	October 10, 1965	17,000
301	April 7, 1948	April 11, 1948	10,000
	August 27, 1949	September 30, 1949	15,000
	July 8, 1954	August 27, 1954	22,000
	August 13, 1954	September 5, 1954	21,000
	November 1, 1956	November 19, 1956	21,000
	April 1, 1957	April 16, 1957	14,000
	April 10, 1959	May 1, 1959	25,000
	April 16, 1959	June 10, 1959	13,000
	April 16, 1959	June 15, 1959	19,000
	June 2, 1965	June 9, 1965	22,000
324	May 15, 1957	September 16, 1957	16,000
331	October 1, 1949	December 1, 1949	500,000
	July 19, 1951	July 24, 1951	12,000
	April 29, 1952	August 15, 1952	560,000
	July 1, 1956	August 5, 1956	500,000
	July 15, 1959	November 8, 1959	519,000
341	December 2, 1953	January 12, 1954	30,000
	March 1, 1965	March 24, 1965	31,000

for each strike month. For industries 242, 271, and 341, instead of omitting all of the necessary observations, in place of the actual value of the output variable which was recorded during the strike month, the value of the output variable during the same month of the previous year was used, multiplied by the ratio of the previous non-strike month's value to the same month of the previous year's value. [For example, if t were a strike month and $t-1$ were a normal month, the value of Y_d used for month t would be Y_{dt-12} (Y_{dt-1}/Y_{dt-13}).] This variable is, in effect, trying to measure what output would have been if the strike had not taken place. In the data appendix these adjustments are presented for the respective industries. For industries 201, 233, and 324, where strikes involving 10,000 or more workers occurred, the strikes did not seem to have a noticeable effect on output, and for these industries no adjustments in the output series were made. Also for the strike ridden industries, 301 and 331, no adjustments in the output series were made, but for these two industries all of the necessary observations were omitted.

Because of the adjustments for shutdowns and strikes, different periods of estimation were used for different industries. The actual period of estimation which was used for each industry is presented in the data appendix. It turned out that the same period of estimation was used for industries 207, 211, 212, 231, 232, 233, and 314, and the same period for industries 311, 332, and 336. For the remaining seven industries the period of estimation was unique to the specific industry.

For the Tires and inner tubes industry, 301, and the Cement industry, 324, monthly data on production were available from the Rubber Manufacturers Association (RMA) and the Bureau of Mines respectively. These data are essentially the same as the FRB data for the industries, since the FRB uses the RMA and Bureau of Mines data to construct the production indices. In this study the RMA data have been used directly for industry 301, and the Bureau of Mines data have been used directly for industry 324. The production data for industry 324 are not available beyond 1964. Whenever the output data were gathered from sources other than the FRB, the monthly figures were converted into average daily rates for the month using the FRB estimate of the number of working days in each month for each industry. This procedure is discussed in detail in the data appendix.

4.3 The results for production workers

4.3.1. The basic results

The basic equation determining the short-run demand for production workers is eq. (3.9), and it is repeated here in the form in which it was estimated:

$$\log M_{2wt} - \log M_{2wt-1} = \alpha_1(\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}) + \alpha_1 \log \overline{H} + \alpha_1\mu t + \sum_{i=1}^{m} \beta_i(\log Y_{dt-i} - \log Y_{dt-i-1}) + \gamma_0(\log Y^e_{dt} - \log Y_{dt-1}) + \sum_{i=1}^{n} \gamma_i(\log Y^e_{dt+i} - \log Y^e_{dt+i-1}). \quad (3.9)'$$

TABLE 4.3

Parameter estimates for eq. (3.9)′ under the expectational hypothesis which gave the better results for each industry

Industry	No. of obs.	$\widehat{\alpha_1 \log H}$	$\hat{\alpha}_1$	$1000\,\widehat{\alpha_1\mu}$	$\hat{\beta}_4$	$\hat{\beta}_3$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\gamma}_0$
201	192	−1.039	−.178	−.077			.074	.067	.265
		(4.54)	(4.54)	(3.90)			(3.96)	(3.60)	(10.16)
207	136	−.874	−.151	.121			.064	.091	.262
		(3.41)	(3.41)	(3.21)			(5.34)	(4.70)	(11.65)
211	136	−.775	−.133	−.050					.086
		(5.79)	(5.80)	(3.20)					(4.43)
212	136	−.636	−.108	−.038				.053	.154
		(4.76)	(4.73)	(1.52)				(4.57)	(7.76)
231	136	−1.056	−.181	.085			.032	.065	.127
		(4.54)	(4.44)	(2.90)			(3.06)	(3.95)	(4.03)
232	136	−.508	−.090	−.062				.021	.118
		(5.50)	(5.57)	(3.40)				(2.93)	(9.59)
233	136	−.048	−.005	.041				.129	.164
		(0.28)	(0.16)	(0.78)				(6.08)	(6.69)
242	154	−.260	−.044	−.011	.060	.105	.146	.150	.218
		(1.42)	(1.41)	(0.55)	(4.57)	(7.35)	(8.62)	(6.87)	(13.65)
271	166	−.258	−.044	−.001					.120
		(2.57)	(2.60)	(0.09)					(7.58)
301	134	−.626	−.108	−.062					.055
		(7.20)	(7.18)	(2.79)					(2.88)
311	170	−1.021	−.174	−.056					.190
		(6.88)	(6.87)	(3.33)					(8.12)
314	136	−.672	−.115	.042					.322
		(2.51)	(2.50)	(2.03)					(10.73)
324	187	−.653	−.110	.060					.224
		(6.37)	(6.34)	(2.44)					(16.50)
331	128	−.209	−.035	.016	.044	.067	.037	.121	.184
		(3.05)	(2.98)	(1.01)	(3.36)	(4.78)	(2.48)	(6.29)	(9.89)
332	170	−.734	−.123	.045					.172
		(8.66)	(8.63)	(2.04)					(8.26)
336	170	−.666	−.113	−.015				.090	.164
		(5.61)	(5.59)	(0.62)				(4.62)	(6.53)
341	191	−.373	−.067	−.060				.038	.182
		(3.60)	(3.62)	(2.38)				(3.88)	(15.32)

t-statistics are in parentheses.

[a] $\hat{\delta}$ is the coefficient estimate of $\log Y_{dt-1} - \log Y_{dt-13}$ under the non-perfect expectational hypothesis.

$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$	$^a\hat{\delta}$	R^2	SE	DW
.171	.119	.138	.159	.087	.074	.039	.665	.0120	1.93
(7.31)	(5.03)	(7.10)	(9.50)	(5.65)	(4.11)	(2.22)			
.182	.125	.080	.034			.056	.855	.0180	2.12
(7.20)	(7.25)	(9.08)	(3.77)			(2.64)			
.024	.043					−.010	.343	.0102	1.92
(1.92)	(4.96)					(0.63)			
							.454	.0159	2.63
.021	.061	.035				−.017	.567	.0194	1.98
(0.97)	(3.95)	(3.57)				(1.02)			
.091	.062	.018					.494	.0107	1.45
(6.77)	(6.28)	(2.54)							
							.512	.0292	1.45
.076	.065						.783	.0126	1.80
(5.27)	(5.20)								
.026	.044	.049	.013	.039			.552	.0048	2.12
(1.68)	(3.37)	(4.26)	(1.34)	(4.06)					
.059	.030	.036					.297	.0142	1.92
(3.37)	(1.83)	(2.29)							
.082	.115	.084	.056	.038			.413	.0115	2.11
(4.66)	(7.51)	(6.05)	(4.52)	(3.50)					
.109	.140	.052				.078	.661	.0143	2.19
(4.28)	(8.04)	(3.80)				(3.34)			
.039	.026	.052	.051			.008	.639	.0177	2.01
(2.40)	(1.60)	(3.36)	(3.42)			(0.47)			
							.790	.0101	1.86
.049	.058	.041	.033				.450	.0167	2.24
(3.46)	(4.57)	(3.45)	(2.82)						
.086	.091	.076	.044	.027			.551	.0175	1.78
(4.83)	(6.00)	(5.79)	(3.42)	(2.13)					
.067	.044	.036	.022				.771	.0180	1.99
(6.02)	(4.64)	(3.87)	(2.45)						

Notice in eq. (3.9)′ that the output variables are the observed FRB variables and not the theoretically more correct Y_{2w} variables. From eq. (3.12) the excess labor variable in eq. (3.9), $\alpha_1(\log M_{2wt-1} - \log M^d_{2wt-1})$, is equal to $\alpha_1(\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}) + \alpha_1 \log \bar{H} + \alpha_1 \mu t$, and this latter expression is the one presented in eq. (3.9)′. The $M^*_{2wt-1}H^*_{2wt-1}$ variable was constructed in the manner described in § 3.6.

Eq. (3.9)′ is, of course, different depending on which expectational hypothesis is assumed. For the perfect expectational hypothesis the actual values of the Y_{dt+i} are used in the equation, and for the other two hypotheses the expectational part of eq. (3.9)′ takes the form presented (for $n = 3$) in eqs. (3.15) and (3.16), with Y_d replacing Y_{2w} in the equations. It was mentioned in § 3.8 that for the two "non-perfect" expectational hypotheses the λ_i coefficients can be identified only if they are all equal. For all hypotheses, however, the γ_i coefficients can be identified.

Each expectational hypothesis for each industry was judged by the goodness of fit of the equation and by the significance of the γ_i coefficients. It turned out that the expectational hypothesis which assumes non-perfect expectations for Y^e_{dt} proved to be substantially inferior in every industry to either of the other two hypotheses, and the results achieved using this hypothesis will not be presented here. These results imply, not too surprisingly, that firms' expectations of the amount of output they are going to produce one month ahead are quite accurate.

The results of estimating eq. (3.9)′ for each of the seventeen industries are presented in table 4.3. For each industry the expectational hypothesis which gave the better results has been used. For the non-perfect expectational hypothesis the coefficient of $\log Y_{dt-1} - \log Y_{dt-13}$ is denoted as δ. The industries for which estimates of δ are given in the table are the industries in which the non-perfect expectational hypothesis proved to be better. In table 4.6 the results of estimating equation (3.9)′ for each industry under the alternative expectational hypothesis to that assumed in table 4.3 are presented, and a comparison of both hypotheses for each industry can be made. Before this comparison is made, however, the results presented in table 4.3 will be discussed.

The results presented in table 4.3 appear to be quite good. For every industry the fit is better than the fit of the basic model of ch. 2 and in most cases substantially so.[1] The coefficient estimates are of the right sign except for two of the estimates of δ, and most of them are highly significant.[2] All of the estimates of the coefficient α_1 of the excess labor variable are negative, and in all but two industries they are highly significant. One of these two industries is industry 242, where the past four changes in output are significant. Without these four variables included in the equation, the estimate of α_1 was significant, but with these variables included it lost its significance. The size of the estimate of α_1 for each industry appears reasonable, with a range of $-.005$ to $-.181$. This implies, other things being equal, an elimination of the amount of excess labor on hand of between about one and twenty percent per month, excluding any effects of the past output change variables. From the results in table 4.3 the amount of excess labor on hand definitely appears to be a significant factor in the determination of short-run changes in the number of workers employed.

With respect to the past output change variables, ignoring the fact that $\log Y_{dt-1} - \log Y_{dt-2}$ is significant in ten of the industries (which may be merely because of the fact that Y_d is used as the output variable instead of Y_{2w}), one or more of the past output change variables are significant in only five of the seventeen industries. It appears, therefore, that the reaction to the amount of excess labor on hand is fairly well specified by the inclusion of the $\log M_{2wt-1} - \log M^d_{2wt-1}$ variable in the equation, and that adding the past output change variables does not add much to the explanation of $\log M_{2wt} - \log M_{2wt-1}$ for most industries.

For every industry the estimate of the coefficient γ_0 of the current output change variable is positive and highly significant. For all but three of the industries one or more of the expected future output change variables are

[1] The results of estimating eq. (2.37) of the basic model of previous studies are presented in tables 2.3 and 2.4.

[2] In estimating eq. (3.9)′ the expected future output change variables were carried forward until they lost their significance, and the past output change variables were carried back until they lost their significance. In five of the industries – 211, 231, 271, 301, and 324 – one or two of the expected future output change variables were not significant, but the ones further out were. In these five cases the insignificant variables were left in.

Because the excess labor variable in eq. (3.9)′ is of the nature of a lagged dependent variable, the classical t-test is only valid asymptotically. Consequently, in the discussion which follows the "t-statistics" will be interpreted somewhat loosely. A coefficient estimate will be said to be "significant" if its t-statistic (the absolute value of the ratio of the coefficient estimate to its standard error) is greater than two. A variable will be said to be "significant" if its coefficient estimate is significant.

TABLE 4.4

Parameter estimates for eq. (3.9)′ with $\alpha_1 \log M_{2wt-1}$ replacing $\alpha_1(\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1})$

Industry	No. of obs.	est. of cnst.	$\hat{\alpha}_1$	coeff. est. of t	$\hat{\beta}_4$	$\hat{\beta}_3$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\gamma}_0$
201	192	.417	−.053	−.003			.099	.112	.160
		(2.35)	(2.34)	(0.20)			(5.22)	(6.83)	(9.93)
207	136	.238	−.036	−.014			.096	.144	.197
		(1.17)	(1.17)	(0.26)			(9.10)	(11.16)	(14.96)
211	136	.185	−.033	.017					.015
		(1.18)	(1.17)	(0.47)					(0.83)
212	136	−.101	.016	.064				.072	.113
		(0.69)	(0.66)	(0.67)				(6.06)	(5.84)
231	136	.377	−.056	.004			.038	.093	.039
		(1.29)	(1.37)	(0.07)			(3.41)	(5.67)	(1.47)
232	136	.292	−.037	.021				.035	.076
		(1.55)	(1.51)	(0.76)				(4.72)	(6.56)
233	136	.880	−.112	.082				.144	.139
		(2.29)	(2.35)	(1.80)				(7.14)	(5.86)
242	154	.210	−.025	−.092	.068	.115	.161	.173	.204
		(1.60)	(1.62)	(1.49)	(5.24)	(8.82)	(12.87)	(13.04)	(13.98)
271	166	.539	−.074	.068					.082
		(3.15)	(3.13)	(2.59)					(7.95)
301	134	.374	−.054	−.085					−.003
		(2.01)	(2.02)	(1.67)					(0.14)
311	170	.056	−.009	−.032					.101
		(0.38)	(0.39)	(0.41)					(4.43)
314	136	1.386	−.179	−.052					.249
		(3.22)	(3.22)	(1.43)					(12.09)
324	187	.087	−.015	−.010					.195
		(0.56)	(0.58)	(0.34)					(12.82)
331	128	−.026	.002	.007	.057	.083	.058	.143	.168
		(0.24)	(0.20)	(0.32)	(4.40)	(6.10)	(4.25)	(7.69)	(8.78)
332	170	−.088	.011	.034					.139
		(0.70)	(0.67)	(1.02)					(5.54)
336	170	−.140	.021	.033				.119	.122
		(1.56)	(1.53)	(1.19)				(5.85)	(4.73)
341	191	.259	−.042	.005				.061	.145
		(1.60)	(1.59)	(0.19)				(7.48)	(17.56)

t-statistics are in parentheses.

[a] $\hat{\delta}$ is the coefficient estimate of $\log Y_{dt-1} - \log Y_{dt-13}$ under the non-perfect expectational hypothesis.

$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$	$^{a}\hat{\delta}$	R^2	SE	DW
.079	.041	.076	.119	.058	.060	.058	.638	.0125	2.01
(5.92)	(2.30)	(4.98)	(7.47)	(3.71)	(3.16)	(3.30)			
.137	.097	.063	.026			.064	.843	.0187	2.20
(5.95)	(5.92)	(7.18)	(2.71)			(2.90)			
−.017	.025					.024	.180	.0114	2.16
(1.40)	(2.60)					(1.43)			
							.363	.0172	2.24
−.031	.033	.025				.018	.506	.0207	1.96
(1.60)	(2.20)	(2.46)				(1.05)			
.061	.041	.005					.382	.0118	1.27
(4.18)	(3.82)	(0.62)							
							.531	.0286	1.40
.064	.056						.784	.0126	1.79
(4.49)	(4.40)								
−.010	.017	.029	.003	.031			.560	.0048	2.11
(1.00)	(1.72)	(2.86)	(0.27)	(3.32)					
.008	−.007	.010					.042	.0166	1.44
(0.38)	(0.35)	(0.52)							
.018	.073	.052	.027	.033			.241	.0130	2.12
(1.03)	(4.38)	(3.30)	(1.90)	(2.56)					
.067	.082	.013				.113	.671	.0141	2.20
(3.14)	(5.48)	(0.93)				(4.67)			
−.024	−.026	.027	.003			.029	.559	.0196	1.94
(1.65)	(1.65)	(1.64)	(0.23)			(1.51)			
							.774	.0105	1.80
.020	.031	.022	.023				.200	.0201	1.77
(1.18)	(2.03)	(1.50)	(1.57)						
.070	.082	.071	.040	.031			.471	.0190	1.77
(3.64)	(4.95)	(4.88)	(2.81)	(2.18)					
.032	.019	.012	.002				.758	.0185	1.94
(3.80)	(2.27)	(1.42)	(0.21)						

also significant. For the fourteen industries in which output expectations are significant, the horizon over which expectations are significant varies from two months for industries 211 and 242 to six months for industry 201.

For the most part in the table the estimates of the γ_i coefficients decrease in size as the expected output changes move further away, which is as expected. The estimate of the coefficient γ_0 of $\log Y^e_{dt} - \log Y_{dt-1}$ is the largest of the output variable coefficient estimates for all of the industries except 301, where the estimate of γ_1 is slightly larger.[1] The over-all results strongly indicate that expected future output changes are significant determinants of the change in the number of production workers employed.

The Durbin–Watson statistics presented in table 4.3 are biased toward two because the excess labor variable ($\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}$) is of the nature of a lagged dependent variable. The bias can be significant, and there may be serial correlation in the model even though the DW statistics in table 4.3 do not indicate so except for industries 212, 232, and 233. This problem will be taken up in detail in ch. 9, but what can be said here regarding the DW statistics is that in general the results show much less evidence of serial correlation than do the results presented in tables 2.3 and 2.4 of estimating the basic model of previous studies. It will be seen in ch. 9 that serial correlation is not a serious problem with respect to the estimates presented in table 4.3.

Because of the fact that the excess labor variable is of the nature of a lagged dependent variable, it might be thought that its significance is due to this fact alone and not because it is measuring excess labor. In econometric time series equations, lagged dependent variables are often significant, and the theoretical reasons for why the lagged dependent variable should be significant are often hard to pin down. The excess labor variable in eq. (3.9)′ is not a simple lagged dependent variable, however, because $\log M^*_{2wt-1}H^*_{2wt-1}$ is substracted from $\log M_{2wt-1}$. This is not a trivial difference, for $M^*_{2wt-1}H^*_{2wt-1}$ has a large short-run variance since it follows fluctuations in output closely. One possible test to use to see whether the excess labor variable is significant merely because $\log M_{2wt-1}$ is significant is to estimate eq. (3.9)′ using $\log M_{2wt-1}$ in place of $\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}$ and see whether $\log M_{2wt-1}$ is significant and whether the fit has been improved by removing the "restriction" from $\log M_{2wt-1}$. Another possible test to use is to estimate eq. (3.9)′ using the lagged dependent variable, log

---

[1] .059 for $\hat{\gamma}_1$ vs. .055 for $\hat{\gamma}_0$.

$M_{2wt-1} - \log M_{2wt-2}$, in place of the excess labor variable and see whether it is significant and whether the fit has been improved.

These two tests were performed, and the results are presented in tables 4.4 and 4.5. For each industry the same equation was estimated as was estimated in table 4.3 (same period of estimation, same number of expectational variables, etc.) except that $\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}$ was replaced by either $\log M_{2wt-1}$ or $\log M_{2wt-1} - \log M_{2wt-2}$. The constant term and the time trend were left in both equations, although there is probably little theoretical justification for including them. There is actually little theoretical justification for either of the equations, and the estimates are meant merely to be used as a test to see whether the excess labor variable is significant merely because it is of the nature of a lagged dependent variable.

In table 4.4 the results of using $\log M_{2wt-1}$ in place of the excess labor variable in eq. (3.9)′ are presented, and in table 4.5 the results of using $\log M_{2wt-1} - \log M_{2wt-2}$ are presented. Looking at the results in table 4.4 first, it is seen that $\log M_{2wt-1}$ is significant in only five of the seventeen industries – 201, 233, 271, 301, and 314. For two of these five industries, 201 and 301, the fit is worse in table 4.4 than in table 4.3 with the excess labor variable included. For industry 301 the difference is substantial ($R^2 = .042$ vs .297), and for this industry none of the output variables is significant in table 4.4. For the other three industries, 233, 271, and 314, the fit in table 4.4 is better than in table 4.3, but only slightly so (for 233 $R^2 =$ .531 vs .512; for 271 $R^2 = .560$ vs .552; for 314 $R^2 = .671$ vs .661). For the twelve industries in which $\log M_{2wt-1}$ is not significant, the fit is worse in table 4.4 than in table 4.3 except for industry 242, where the fits are essentially the same. (Remember that the excess labor variable was not significant for industries 242 because of the inclusion of the past four output change variables.) For most of the twelve industries except 242, the fits are not only worse in table 4.4 than in table 4.3 but are substantially worse, and many of the output variables are either not significant or less significant in table 4.4 than they were in table 4.3. The over-all results definitely indicate that the excess labor variable, $\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}$, is not significant merely because $\log M_{2wt-1}$ is: $\log M_{2wt-1}$ is not in general significant by itself and using $\log M_{2wt-1}$ in place of $\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}$ has considerably worsened the over-all results.

Turning now to the results in table 4.5, it is seen that $\log M_{2wt-1} - \log M_{2wt-2}$ is significant in only six of the industries – 231, 271, 301, 311, 324, and 332. For two of these industries the coefficient estimate of $\log M_{2wt-1} - \log M_{2wt-2}$ is negative, and for the other four the estimate is positive. For

TABLE 4.5

Parameter estimates for eq. (3.9)′ with $\alpha_1(\log M_{2wt-1} - \log M_{2wt-2})$ replacing $\alpha_1(\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1})$

Industry	No. of obs.	est. of cnst.	$\hat{\alpha}_1$	coeff. est. of t	$\hat{\beta}_4$	$\hat{\beta}_3$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\gamma}_0$
201	192	.001	−.004	−.017			.096	.111	.169
		(0.57)	(0.06)	(1.08)			(4.55)	(5.73)	(10.20)
207	136	−.001	−.076	.036			.109	.153	.219
		(0.25)	(1.21)	(1.21)			(6.15)	(9.64)	(10.10)
211	136	.001	.106	−.019					.026
		(0.58)	(1.48)	(1.18)					(1.43)
212	136	−.003	−.091	.002				.077	.103
		(0.98)	(1.96)	(0.06)				(6.39)	(5.42)
231	136	−.025	−.203	.072			.052	.128	.039
		(5.60)	(2.64)	(2.37)			(4.21)	(6.15)	(1.50)
232	136	.007	−.016	−0.12				.036	.080
		(2.47)	(0.26)	(0.69)				(4.68)	(6.99)
233	136	−.020	.113	.050				.103	.180
		(3.43)	(1.82)	(1.14)				(4.16)	(7.51)
242	154	−.003	.024	.005	.062	.109	.158	.170	.207
		(1.29)	(0.29)	(0.26)	(4.06)	(6.04)	(9.15)	(8.38)	(12.98)
271	166	.003	−.168	−.103					.080
		(4.09)	(3.03)	(2.13)					(7.67)
301	134	−.002	.216	.010					.001
		(0.64)	(2.62)	(0.41)					(0.04)
311	170	−.002	.164	−.002					.093
		(0.84)	(2.49)	(0.13)					(4.21)
314	136	−.003	.154	.045					.268
		(1.01)	(1.83)	(2.15)					(13.15)
324	187	−.002	.145	−.003					.184
		(0.66)	(2.51)	(0.14)					(12.22)
331	128	−.005	−.065	.004	.063	.086	.065	.155	.170
		(2.03)	(0.77)	(0.25)	(4.38)	(6.05)	(3.99)	(6.37)	(9.13)
332	170	−.003	.200	.019					.127
		(0.85)	(2.92)	(0.74)					(5.25)
336	170	−.003	.132	.010				.105	.107
		(0.76)	(1.96)	(0.42)				(4.87)	(4.05)
341	191	.002	.003	−.018				.059	.149
		(0.68)	(0.04)	(0.78)				(4.68)	(18.17)

t-statistics are in parentheses.

[a] $\hat{\delta}$ is the coefficient estimate of $\log Y_{dt-1} - \log Y_{dt-13}$ under the non-perfect expectational hypothesis.

$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$	[a]$\hat{\delta}$	R^2	SE	DW
.082	.046	.080	.126	.065	.069	.061	.627	.0127	2.04
(6.06)	(2.45)	(4.84)	(7.66)	(4.16)	(3.65)	(3.43)			
.152	.106	.070	.033			.070	.844	.0187	2.09
(6.09)	(6.22)	(8.12)	(3.36)			(3.14)			
−.013	.032					.024	.185	.0114	2.54
(1.10)	(3.32)					(1.48)			
							.379	.0170	1.74
−.037	.028	.018				.014	.525	.0203	1.87
(2.07)	(1.93)	(1.67)				(0.86)			
.067	.045	.006					.371	.0119	1.30
(4.69)	(4.27)	(0.80)							
							.524	.0288	1.58
.069	.062						.780	.0127	1.88
(4.82)	(4.98)								
−.003	.024	.035	.008	.053			.559	.0048	1.88
(0.29)	(2.52)	(3.42)	(0.85)	(4.82)					
.021	.001	.019					.062	.0164	2.02
(1.09)	(0.08)	(1.05)							
.011	.067	.052	.027	.030			.269	.0128	2.49
(0.61)	(4.19)	(3.53)	(2.06)	(2.38)					
.080	.125	.055				.080	.653	.0144	2.28
(3.67)	(8.23)	(3.37)				(3.40)			
−.014	−.018	.021	.017			.023	.573	.0192	2.26
(0.93)	(1.18)	(1.30)	(1.12)			(1.22)			
							.775	.0104	1.64
.009	.026	.015	.019				.238	.0196	2.39
(0.56)	(1.83)	(1.10)	(1.38)						
.058	.070	.063	.032	.026			.476	.0189	2.03
(3.02)	(4.17)	(4.43)	(2.34)	(1.86)					
.038	.023	.017	.006				.755	.0187	1.98
(4.77)	(2.94)	(2.14)	(0.78)						

industry 271 the fit is slightly better in table 4.5 than in table 4.3 ($R^2 = .559$ vs .552), but for the other five industries the fit is substantially worse in table 4.5 (for 231 $R^2 = .525$ vs .567; for 301 $R^2 = .062$ vs .297; for 311 $R^2 = .269$ vs .413; for 324 $R^2 = .573$ vs .639; for 332 $R^2 = .238$ vs .450). For the eleven industries in which $\log M_{2wt-1} - \log M_{2wt-2}$ is not significant, the fit is worse (and in many cases substantially worse) in table 4.5 than in table 4.3 except for industry 233, where the fit is slightly better in table 4.5 ($R^2 = .524$ vs .512). The output variables in table 4.5 are in general much less significant than in table 4.3. The use of $\log M_{2wt-1} - \log M_{2wt-2}$ in place of the excess labor variable, $\log M_{2wt-1} - \log M^*_{2wt-1}H^*_{2wt-1}$, therefore, has considerably worsened the over-all results, and it seems safe to conclude from the results in tables 4.4 and 4.5 that the excess labor variable is significant in its own right and not because it is of the nature of a lagged dependent variable. Some further results will be presented in ch. 7 which indicate that the excess labor variable is also significant in the equation determining the change in the number of hours paid-for per worker, and these results are an independent confirmation of the hypothesis that the amount of excess labor on hand affects short-run employment decisions.

In a monthly model such as this one there is the possibility that the behavior of $\log M_{2wt} - \log M_{2wt-1}$ is significantly different during one specific month of the year than during the other eleven months. To the extent that the model is well specified this should not be the case, but there may be factors influencing $\log M_{2wt} - \log M_{2wt-1}$ in a systematic way during the same month each year which have not been taken into account in the model. One possible test to use to test whether this is true is the F-test, testing the hypothesis that the coefficients for one specific month of the year are the same as the coefficients for all of the other months. A cruder test was in fact performed in this study. For each industry for each month the number of positive and negative residuals was calculated to see if there was a systematic tendency for the estimated equation to underpredict or overpredict for a specific month. Assuming that the probability of any one residual being negative is one-half, the hypothesis that the residuals for any one month come from a binomial population (with $p = 1/2$) was rejected (at the five-percent confidence level) in 37 of the 162 cases, or in about 23 percent of the cases.[1]

[1] It should be emphasized that this test is crude and that the results of the test should be interpreted as indicating only general tendencies.

Six of these cases occurred for the June–May period (where the model underpredicted) and seven for the October–September period (where the model overpredicted). The student influx in early June and outflow in middle September probably account for this situation. Four of the cases occurred for the December–November period (where the model underpredicted), and these are probably accounted for by the fact that for December the average daily rate of output for the month is likely to be much less than the rate during the second week. Five of the cases occurred for the March–February period (where the model overpredicted), and this may be accounted for by the fact that for March the average daily rate of output for the month may be greater than the rate during the second week if the spring upturn begins during the last half of March. The other fifteen cases were about evenly distributed over the remaining months and showed no systematic tendency to underpredict or overpredict for a particular month.

For the 77 percent of the cases where the hypothesis was not rejected, the residuals appeared to be fairly random. The general conclusion of this test is that while there are some systematic tendencies by month which the model fails to account for, some of which can be explained by faulty data and some by student inflows and outflows, the model in general seems to do reasonably well.

In summary, then, the results appear to be quite good. The amount of excess labor on hand definitely appears to be a significant factor in the determination of the change in the number of workers employed, and the time stream of expected future changes in output also appears to be a significant factor. The coefficient estimates are of the right sign and in most cases are highly significant. For every industry the fit is better than the fit of the basic model of previous studies, and for most industries it is substantially better. There seems to be only a few monthly systematic tendencies which the model has not explained, and it will be seen later that in general the residuals of the equations are not serially correlated.

4.3.2. A comparison of the expectational hypotheses

In table 4.6 the results are presented of estimating eq. (3.9)′ for each industry under the alternative (and inferior) expectational hypothesis to that assumed

TABLE 4.6

Parameter estimates for eq. (3.9)′ under the alternative (and inferior) expectational hypothesis to that used in table 4.3 for each industry

Industry	No. of obs.	$\widehat{\alpha_1 \log \bar{H}}$	$\hat{\alpha}_1$	$1000\,\widehat{\alpha_1 \mu}$	$\hat{\beta}_4$	$\hat{\beta}_3$	$\hat{\beta}_2$	$\hat{\beta}_1$	$\hat{\gamma}_0$
201	192	−.430	−.073	−.026			.064	.103	.201
		(2.52)	(2.52)	(1.46)			(3.23)	(5.68)	(8.74)
207	136	−.775	−.134	.074			.064	.109	.245
		(2.80)	(2.79)	(2.03)			(5.03)	(5.16)	(10.00)
211	136	−.674	−.115	−.047			.079	.012	.039
		(5.82)	(5.83)	(3.00)			(4.21)	(0.94)	(4.56)
231	136	−.805	−.136	.084			.031	.066	.094
		(6.27)	(6.04)	(2.88)			(2.90)	(4.05)	(4.08)
232	136	−.385	−.068	−.031				.030	.088
		(2.88)	(2.89)	(1.47)				(3.57)	(5.25)
242	154	−.042	−.006	.006	.066	.111	.156	.166	.222
		(0.23)	(0.21)	(0.31)	(4.78)	(7.39)	(8.77)	(7.37)	(13.14)
271	166	−.304	−.052	.005					.125
		(2.11)	(2.12)	(0.55)					(5.51)
311	170	−.911	−.155	−.044					.164
		(3.97)	(3.96)	(2.23)					(4.60)
314	136	−.599	−.102	.043					.300
		(2.86)	(2.84)	(2.02)					(11.58)
324	187	−.668	−.112	.066					.221
		(6.28)	(6.29)	(2.53)					(13.99)
332	170	−.367	−.061	.035					.133
		(2.51)	(2.49)	(1.42)					(4.58)
336	170	−.228	−.038	−.001				.132	.126
		(1.15)	(1.12)	(0.03)				(6.05)	(3.54)
341	191	−.227	−.041	−.045				.038	.173
		(2.15)	(2.16)	(1.65)				(3.69)	(14.24)
301	[b]99	−.591	−.102	−.054					.059
		(4.41)	(4.36)	(1.68)					(2.23)
	[b]99	−.351	−.060	−.010					−.007
		(1.80)	(1.78)	(0.31)					(0.22)

t-statistics are in parentheses.

[a] $\hat{\delta}$ is the coefficient estimate of $\log Y_{at-1} - \log Y_{at-13}$ under the non-perfect expectational hypothesis.

[b] Different period of estimation used here than in table 4.3.

$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\gamma}_3$	$\hat{\gamma}_4$	$\hat{\gamma}_5$	$\hat{\gamma}_6$	$^a\hat{\delta}$	R^2	SE	DW
.106	.046	.124	.140	.058	.042		.607	.0130	1.72
(5.07)	(2.05)	(6.70)	(8.09)	(3.42)	(2.08)				
.161	.126	.075	.034				.830	.0195	1.90
(5.86)	(7.06)	(8.10)	(3.46)						
							.338	.0102	1.93
−.019	.037	.020					.562	.0194	2.13
(0.99)	(2.60)	(2.03)							
.021	.023	.000				.013	.314	.0124	1.54
(1.46)	(2.10)	(0.03)				(1.10)			
.052	.064					.034	.770	.0131	1.88
(3.74)	(4.94)					(2.30)			
.035	.040	.048	.019	.030		.024	.517	.0050	2.13
(1.75)	(2.23)	(3.33)	(1.59)	(2.85)		(2.03)			
.028	.068	.049	.034	.023		−.010	.271	.0128	1.92
(1.25)	(3.42)	(2.84)	(2.17)	(1.67)		(0.63)			
.114	.133	.040					.637	.0147	2.05
(4.12)	(8.34)	(2.98)							
.042	.026	.046	.053				.630	.0179	2.03
(1.81)	(1.31)	(2.37)	(2.86)						
−.027	.001	−.000	.001			−.006	.371	.0179	2.06
(1.74)	(0.06)	(0.03)	(0.08)			(0.52)			
−.032	.023	.032	.019	.017		−.019	.447	.0195	1.83
(1.73)	(1.26)	(1.87)	(1.17)	(1.07)		(1.24)			
.047	.043	.025	.010			.038	.767	.0183	1.97
(4.19)	(4.47)	(2.67)	(1.03)			(2.96)			
.067	.049	.037					.212	.0151	2.06
(2.93)	(2.37)	(1.85)							
−.032	−.036	−.041				−.034	.180	.0155	2.14
(1.56)	(1.41)	(1.74)				(1.84)			

in table 4.3.[1] For each industry the same expectational time horizon (size of n) was used as proved to be significant for the estimates in table 4.3. For industries 212, 233, and 331 none of the expected future output change variables was significant under either expectational hypothesis, and thus no comparison for these industries is needed.[2]

For six industries – 201, 207, 211, 231, 314, and 324 – the non-perfect expectational hypothesis is superior. Examining the results for these industries in the two tables reveals that the perfect expectational hypothesis works almost as well in all six industries. For the perfect expectational hypothesis the estimates of the γ_i coefficients are nearly as significant as for the other hypothesis, and the fits are nearly as good. Industry 201 shows the most difference between the two hypotheses, but even in this case the perfect expectational hypothesis does not perform badly.

In three of the six industries where the non-perfect expectational hypothesis gives the better results, the estimates of the coefficient δ of $\log Y_{dt-1} - \log Y_{dt-13}$ are not significant, which, under the assumption that all of the λ_i coefficients are equal, implies that the rate of output in a specific future month is expected to be equal to what the rate of output was during the same month of the preceeding year. Expectations in this case are static.

For the remaining eight industries – 232, 242, 271, 301, 311, 332, 336, and 341 – the perfect expectational hypothesis is superior. Examining the results for these industries in the two tables reveals that the non-perfect expectational hypothesis works almost as well for industries 242, 271, and 341. For the five industries, 232, 301, 311, 332, and 336, however, the non-perfect expectational hypothesis yields substantially inferior results than

[1] As mentioned in § 4.2, for industry 301 none of the strike-period observations for output was changed and all of the necessary observations were omitted. Since the non-perfect expectational hypothesis involved longer lags, it was necessary to omit more observations under this hypothesis than under the perfect expectational hypothesis. To make the results for this industry comparable, therefore, the equation for the perfect expectational hypothesis was re-estimated using the same period of estimation as was used to estimate the equation for the non-perfect expectational hypothesis. In table 4.6 both of these results are presented, and a comparison of the expectational hypotheses for this industry is made using these results.

[2] For industry 331 none of the strike-period observations for output was changed, and when estimating eq. (3.9)′ for this industry under the non-perfect expectational hypothesis, a different period of estimation (in which more observations were omitted) was used than was used in table 4.3. None of the expected future output change variables was significant for either period of estimation, and so the larger period of estimation was used for the estimates in table 4.3.

does the perfect expectational hypothesis, both on grounds of goodness of fit and significance of the γ_i coefficient estimates. The fits are much worse[1] and most of the estimates of the γ_i coefficients are not significant.

It is thus quite evident that if one had to choose between the two hypotheses, he would choose the perfect expectational hypothesis as giving the better results. In all fourteen industries for which future output expectations are significant at all, the perfect expectational hypothesis gives good results and results better than the other hypothesis in eight of the fourteen industries. The non-perfect expectational hypothesis, on the other hand, gives good results for only nine of the fourteen industries. It was mentioned in § 3.8 that the perfect expectational hypothesis may be a better approximation of reality if firms can in fact forecast more accurately than the naive, non-perfect expectational hypothesis says they can or if firms schedule production in advance and do not deviate much from this schedule even if demand conditions change. In this study and in the work which follows, an absolute choice was not in fact made between the two hypotheses, and the hypothesis which gave the better results for a particular industry was assumed to be true for that industry.

One other comparison was actually made of the two expectational hypotheses. In addition to the assumption that $Y_t^e = Y_t$, the following additional assumptions were made:

$$\log Y^e_{dt+1} - \log Y_{dt} = \zeta_1(\log Y_{dt-11} - \log Y_{dt}) + (1 - \zeta_1)(\log Y_{dt+1} - \log Y_{dt}), \tag{4.1a}$$

$$\log Y^e_{dt+2} - \log Y^e_{dt+1} = \zeta_2(\log Y_{dt-10} - \log Y_{dt-11}) + (1 - \zeta_2)(\log Y_{dt+2} - \log Y_{dt+1}), \tag{4.1b}$$

$$\log Y^e_{dt+3} - \log Y^e_{dt+2} = \zeta_3(\log Y_{dt-9} - \log Y_{dt-10}) + (1 - \zeta_3)(\log Y_{dt+3} - \log Y_{dt+2}), \tag{4.1c}$$

and so on. These assumptions are in a sense a weighted average of the two expectational hypotheses.[2] For the perfect expectational hypothesis all of

---

[1] For 232, R^2 = .314 vs .494.
For 301, R^2 = .180 vs .212.
For 311, R^2 = .271 vs .413.
For 332, R^2 = .371 vs .450.
For 336, R^2 = .447 vs .551.

[2] This type of assumption is similar to that made by Lovell in his study of inventory investment. See LOVELL (1961, p. 305).

TABLE 4.7

Parameter estimates for eq. (3.9)′ under the assumptions made in eq. (4.1) about expectations. Estimates presented only for the coefficients of the expectational variables

Industry	No. of obs.	$\widehat{\gamma_1\zeta_1}$	$\widehat{\gamma_2\zeta_2}$	$\widehat{\gamma_3\zeta_3}$	$\widehat{\gamma_4\zeta_4}$	$\widehat{\gamma_5\zeta_5}$	$\widehat{\gamma_1(1-\zeta_1)}$	$\widehat{\gamma_2(1-\zeta_2)}$	$\widehat{\gamma_3(1-\zeta_3)}$	$\widehat{\gamma_4(1-\zeta_4)}$	$\widehat{\gamma_5(1-\zeta_5)}$	R^2	SE	DW
207	136	.152	.089	.058	.029		.051	.055	.021	.007		.851	.0185	2.01
		(4.17)	(2.76)	(2.07)	(1.30)		(1.30)	(1.65)	(0.76)	(0.30)				
232	136	.002	.014	.015			.089	.050	.005			.507	.0106	1.46
		(0.19)	(1.28)	(1.46)			(6.45)	(4.05)	(0.52)					
242	154	.011	.032				.066	.041				.788	.0126	1.84
		(0.97)	(1.64)				(4.11)	(2.06)						
311	170	.018	.041	.028	.021	−.023	.089	.095	.073	.047	.061	.435	.0114	2.12
		(1.11)	(1.82)	(1.26)	(1.02)	(1.14)	(4.89)	(4.73)	(3.77)	(2.51)	(3.31)			
314	136	.046	.098	.047			.097	.060	.008			.666	.0143	2.12
		(2.08)	(3.29)	(1.65)			(3.49)	(2.09)	(0.29)					

Implied values of the ζ_i coefficients

Industry	ζ_1	ζ_2	ζ_3	ζ_4	ζ_5
207	.749	.617	.734	.806	
232	.022	.219	.750		
242	.143	.438			
311	.168	.301	.277	.309	−.605
314	.322	.620	.855		

t-statistics are in parentheses.

the ζ_i's are zero, and for the non-perfect expectational hypothesis all of the ζ_i's are one and the $\log Y_{dt-1} - \log Y_{dt-13}$ variable is added.

Eq. (3.9)′ was then estimated under the assumptions made in (4.1) for each of the fourteen industries for which future output expectations are significant. Again, the same size of n was used for each industry as proved to be significant in table 4.3. For five industries – 271, 301, 332, 336, and 341 – it was obvious that the ζ_i coefficients were not significantly different from zero. The output variables representing the perfect expectational hypothesis completely dominated the output variables representing the non-perfect expectational hypothesis. For four industries – 201, 211, 231, and 324 – it was obvious that the $(1 - \zeta_i)$ coefficients were not significantly different from zero. For these industries the output variables representing the non-perfect expectational hypothesis completely dominated the output variables representing the perfect expectational hypothesis. These results are consistent with the results reported above of estimating eq. (3.9)′ under each expectational hypothesis separately. As was seen in tables 4.3 and 4.6, for industries 271, 301, 332, 336, and 341 the perfect expectational hypothesis gave the better results, and for industries 201, 211, 231, and 324 the non-perfect expectational hypothesis performed better.

For the five remaining industries – 207, 232, 241, 311, and 314 – one hypothesis likewise appeared to dominate the other, but since this domination was not quite as evident for these industries, it is worthwhile to examine the results more closely. In table 4.7 the results of estimating eq. (3.9)′ under the assumptions made in (4.1) are presented for these five industries. The estimates are given only for the coefficients of the expectational variables, as the other coefficient estimates were little changed. Also presented in table 4.7 are the derived values of the ζ_i coefficients.

For industry 207 the estimates of the coefficients of the output variables representing the non-perfect expectational hypothesis are larger and more significant than the estimates of the coefficients of the other output variables (which are small and not significant). The size of ζ_i coefficients ranges between .617 and .806. For industries 232, 242, and 311 the estimates of the coefficients of the output variables representing the perfect expectational hypothesis are in general larger and more significant, but there does appear to be a tendency for the estimates of the coefficients of the output variables representing the non-perfect expectational hypothesis to become larger and more significant relative to the estimates of the coefficients of the other output variables as the period for which the expectation is made moves further into the future. In other words, there seems to be a tendency for ζ_i

to increase as i increases. This is definitely true for industries 232 and 242, and slightly true for 311, except for the last coefficient, ζ_5, which is in fact negative. Industry 314 gives the best results for assumption (4.1). Except for the last period, the estimates of the coefficients of the output variables representing each hypothesis are significant. There is also clear evidence that ζ_i increases as i increases for this industry.

This slight over-all evidence that ζ_i increases as i increases is consistent with theoretical notions, as one would expect that as the periods for which the expectations are made move further into the future, there will be less ability to predict accurately and more of a tendency to rely on past behavior. The results in general indicate, however, that the "weighted average" assumptions made in (4.1) are not an improvement over either the perfect expectational hypothesis or the non-perfect expectational hypothesis considered separately. The fits are little changed over those in table 4.3, and in general one set of output variables dominates the other set.

4.4 Summary

The results presented in this chapter appear to be an important confirmation of the model of the short-run demand for workers which was developed in ch. 3. Using monthly data for seventeen three-digit United States manufacturing industries, the basic equation of the model [eq. (3.9)′] was estimated under the perfect and non-perfect expectational hypotheses. For every industry the fit of eq. (3.9)′ was better than the fit of eq. (2.37) of the basic model of previous studies and for most industries was considerably better. The excess labor variable in eq. (3.9)′ was significant in all but two industries, one in which the past four output change variables were significant. In all of the industries the current output change variable was significant, and for all but three of the industries at least two other expected future output change variables were also significant. The results indicate, therefore, that both the amount of excess labor on hand and the time stream of expected future output changes are significant determinants of the change in the number of production workers employed.

Regarding the expectational hypotheses, the perfect expectational hypothesis gave somewhat better over-all results than the non-perfect expectational hypothesis, but the latter gave slightly better results for six of the fourteen industries where future output expectations were significant. The decision was made to use the non-perfect expectational hypothesis for these six industries and the perfect expectational hypothesis for the eight others.