

NBER WORKING PAPER SERIES

/HETEROGENEITY IN PANEL DATA:/ ARE THERE
STABLE PRODUCTION FUNCTIONS?

Jacques Mairesse
Zvi Griliches

Working Paper No. 2619

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 1988

The research supported here is part of the NBER's research program in Productivity. Any opinions expressed are those of the authors and not those of the National Bureau of Economic Research.

HETEROGENEITY IN PANEL DATA: ARE THERE STABLE PRODUCTION FUNCTIONS?

Jacques Mairesse and Zvi Griliches

Errata

Due to the vagaries of the word processing system and the travel of the authors several errors crept into this version. In particular, Section 6, "Swamy Estimates," should be disregarded. It is superseded by the comparably titled appendix. Also, as several references are not included, a complete listing follows:

References

- Amemiya, T. 1977. "A Note on a Heteroskedastic Model," Journal of Econometrics 5: 295-299.
- Carter, R.L. and M.C.K. Yang. 1986. "Large Sample Inference in Random Coefficient Regression Models," Communications in Statistics, Part A, Theory and Methods, 15(8): 2507-2525.
- Chamberlain, G. 1984. "Panel Data," in Handbook of Econometrics, Vol. II, ed. by Z. Griliches and M. Intriligator, 1247-1318. Amsterdam: North-Holland.
- Cox, D. 1983. "Some Remarks on Overdispersion," Biometrika 70.
- Cragg, J.G. 1987. "A Bilinear Model for Heteroskedastic Panel Data," University of British Columbia, Department of Economics Discussion Paper No. 87-23, Vancouver, Canada.
- Dutta, J. and H.L. Leon. 1986. "Testing for Heterogeneous Parameters in a Least Squares Framework," First Boston Working Paper Series 86-32, Columbia University, Graduate School of Business.
- Goldfeld, S.R. and R.M. Quandt. 1972. Non Linear Methods in Econometrics. Amsterdam: North-Holland.

- Gourieroux, C., A. Monfort and A. Trognon. 1984. "Pseudo Maximum Likelihood Methods," Econometrica 52: 681-720.
- Griliches, Z. and J. Mairesse. 1983. "Comparing Productivity Growth: An Exploration of French and U.S. Industrial and Firm Data," European Economic Review 21: 89-119.
- _____. 1984. "Productivity and R&D at the firm level," in R-D, Patents and Productivity, ed. by Z. Griliches, 339-374. The University of Chicago Press.
- _____. 1985. "R&D and Productivity Growth: Comparing Japanese and U.S. Manufacturing firms," National Bureau of Economic Research Working Paper No. 1778. To appear in an "Income and Wealth" conference volume.
- Grunfeld, Y. and Z. Griliches. 1960. "Is Aggregation Necessarily Bad?," Review of Economics and Statistics, :1-13.
- Hausman, J.A. 1978. "Specification Tests in Econometrics," Econometrica 46: 1251-1273.
- Hildreth, C. and J.P. Houck. 1968. "Some Estimates for a Linear Model with Random Coefficients," Journal of the American Association 63: 584-595.
- Hsiao, C. 1986. Analysis of Panel Data. Cambridge: Cambridge University Press.
- Leamer, E.E. 1978. Specification Searches: Ad hoc Inference with Nonexperimental Data. New York: John Wiley and Sons.
- MaCurdy, T. 1981. "Multiple Time-Series Models Applied to Panel Data," National Bureau of Economic Research Working Paper No. 646.

- _____. 1985. "A Guide to Applying Time Series Models to Panel Data," Stanford University, unpublished paper.
- Mairesse, J. 1987. "Les lois de la production ne sont plus ce qu'elle etaient: une introduction a l'econometrice des panels," La Revue Economique, (1): 225-271.
- Mundlak, Y. 1978. "On the Pooling of Time Series and Cross-Section Data," Econometrica 46: 69-85.
- Scheffe, H. 1959. The Analysis of Variance. New York: John Wiley and Sons.
- Solon, G. 1984. "Estimating Autocorrelation in Fixed Effects Models," National Bureau of Economic Research Technical Working Paper No. 32.
- Swamy, P.A.V.B. 1970. "Efficient Inference in a Random Coefficient Model," Econometrica 38: 311-23.
- _____. 1971. Statistical Inference in Random Coefficient Regression Models. Berlin: Springer-Verlag.
- White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimation and a Direct Test for Heteroskedasticity," Econometrica 48: 817-838.
- _____. 1982. "Maximum Likelihood Estimators of Misspecified Models," Econometrica 50(1): 1-25.

208550

NBER Working Paper #2619
June 1988

Heterogeneity In Panel Data: Are There
Stable Production Functions?

ABSTRACT

We estimate separate production functions for approximately 450 manufacturing firms each in France and the United States and for 850 manufacturing firms in Japan, covering the 13 year period 1967-1979, and focus on the wide dispersion in the estimated slope coefficients in all three countries. The main question asked is: "Is this dispersion real?" Could it be just a reflection of sampling variability or is it an indication of real heterogeneity? We estimate the "true" dispersion using three different approaches: Maximum Likelihood, regressions of squares and cross-products of residuals, and Swamy's "residual" method, and try to interpret the somewhat different answers which emerge. In particular, we investigate the "reality" of the estimated heterogeneity by looking at its stability over time and by relating it to differences in capital shares and the industrial structure. We conclude that the observed heterogeneity is not "real." It is caused by some non-stable misspecification of our simple model, implying that we are unlikely to discern different but stable individual production relations in samples of this size which contain only a limited number of the economically relevant variables.

Jacques Mairesse
Institut National de la
Statistique et des Etudes
Economiques
18, Boulevard Adolphe-Pinard
75675 Paris Cedex 14
FRANCE

Zvi Griliches
National Bureau of Economic Research
1050 Massachusetts Avenue
Cambridge, MA 02138

November 1987
Revised May 1988

HETEROGENEITY IN PANEL DATA: ARE THERE STABLE PRODUCTION FUNCTIONS?

Jacques Mairesse and Zvi Griliches*

1. INTRODUCTION AND MOTIVATION

Our paper is about heterogeneity in production relations at the micro level and a quest for an interpretation of it. We have been studying, for a number of years, the relationship between output and inputs, between labor productivity and the capital-labor ratio, and between technological change and investment in R&D. We have worked primarily with firm level data in France, the U.S. and Japan and used the production function framework, and especially the Cobb-Douglas form, as an organizing device for our analysis.¹ In this work we were struck both by the surprising amount of variability in the basic economic ratios at the micro level and by a number of "anomalies" in the results of conventional analyses of such data. Since the usual approaches impose a common specification with constant parameters across the different units, it is at least possible that some of the apparent anomalies in our results could be due to the neglect of the parameter heterogeneity.

We ask, in this paper, if the observed heterogeneity is "real" in some sense and if it makes a difference for our major conclusions. It will turn out, however, that it is much more difficult to provide a clear answer to the first part of this question and we will, therefore, make only very

little progress on the second one. To illustrate the problem and to describe some of the major aspects of our data we shall focus, primarily, on analyzing estimates of the capital coefficient in a constant returns Cobb-Douglas production function and their potential heterogeneity across different firms. We shall make a number of simplifying assumptions to make the problem tractable and to help us to communicate our major results. In particular, we shall abstract from various issues of simultaneity and other obvious specification errors and conduct analysis under the assumption of constant returns to scale (CRS).²

The story begins with our data: We have collected, constructed, cleaned, and analyzed annual data on the output and inputs of about 450 manufacturing firms each in France and the U.S. and 850 firms in Japan, for the years 1967 through 1979. We imposed common production function coefficients across firms, and often also across industries within manufacturing, in order to provide us with a framework for an analysis of the sources of productivity growth in different countries and also to help us in exploring the contribution of R&D. We estimated simple production functions as a first step in such analyses. Continued experience with the data has convinced us that most of the encountered problems cannot be cured by fancier functional forms. We did not work with "complete systems" of production relations. These require the availability of relevant individual firm input and output prices at the micro level, something which is not within the feasible set at the moment.

Table 1 shows the result of fitting a standard CRS Cobb-Douglas production function to various dimensions of the data. We present three estimates of the capital coefficient (elasticity) in two settings: levels

and first differences. The first column gives the "total" results, the result of pooling all the data and imposing a common intercept and common slope on all the observations. The data sets are then divided into their "between" and "within" components, the first being equivalent to estimating the equation on average, over 13 years, for each firm, while the second allows each firm to have its own intercept and gets its estimate from the "within-firm" time series variations. It allows thereby for the presence of correlated individual firm effects in the sense of Mundlak (1978) and Chamberlain (1984). It should be noted also, that all the equations contain a set of fourteen industry specific year dummy variables, allowing thereby for flexible and industry specific time trends. The first difference estimators are conceptually close to the "within" estimators in the sense that they also allow for the presence of individual firm constants by differencing them out of the data. Their "between" component is equivalent to the "longest" difference: i.e., it estimates the coefficients from the average growth rate of the variables over the whole 13 year period. The "within" first differences estimator allows also for individual firm differences in trends.

The overall results are not particularly surprising. Total and between estimates are close together [the bulk of the variance in the data is between firms and accounts for about 88 percent of the total variance in the logarithm of labor productivity in all three countries] and are higher, especially for France and Japan, than the within estimates. Japan tends to have the highest capital elasticity (except in the "within" first differences). Only in the U.S. data are the "between" and "within" differences minor. These differences are even more striking, also in the

U.S. data, if one does not impose the constant returns to scale assumption (see Mairesse, 1988). These are the kind of "anomalies" that have preoccupied us for some time.

Now the fact that "within" estimates of economic coefficients tend to be lower than "between" estimates is not new and has been much commented on in the past. The very simple model that underlies these estimates may be subject to many misspecifications which could account for such discrepancies. An incomplete list would include the assumption of constant returns to scale, the use of gross rather than net output measures and the omission of measures of material use, R&D capital, and other aspects of quality of labor and capital; the lack of measures of input utilization and the associated simultaneity problems and short-run versus long-run productivity movements issues, and the possibility of serious measurement error in most of our variables. Some of these issues have already been explored in other contexts and on other occasions. Several of the others, especially issues of capacity utilization and measurement error we plan to pursue further in the near future. Here we want to explore the underlying heterogeneity in firm responses as another possible source of the observed anomalies in our results.

The reason why we thought of this as a promising line of attack is that if firms really differ in their response coefficients, estimates based on different cuts of the data would represent a different weighting of the underlying distribution of parameters. For example, "between" estimates of production coefficients would weight the individual coefficients in proportion to the square distance between their input ratios and the economy or industry wide averages while "within" estimates of the same

coefficients would weight the same individual parameters by the relative variance of their input ratios over time. The two weight structures are very different and if there were some correlation between such weights and the underlying true parameters, the resulting average parameters could differ quite a bit from each other even if the underlying parameter distribution was stable in some sense. A negative correlation of the individual parameters with the "within" weights could account for the finding that the "within" estimates are, in general, smaller than the "between." Such a correlation could arise if firms whose environments are changing more rapidly and whose capital-labor ratios fluctuate more have managed to acquire technologies with less responsive, smaller capital coefficients.³

To see whether this could indeed provide at least a partial explanation for our observed results we embarked on the task of estimating a separate production function for each of our firms in all three countries and summarizing the massive results in some comprehensible fashion. It was clear from the beginning that with only 13 observations for each firm we could estimate only rather simple, few parameter relationships, and that the individual results were unlikely to be very precise. We were hoping to make up what we were losing in precision at the individual level by a clearer picture of the distribution of the parameters across firms. We were not prepared, however, for the amount of heterogeneity that we encountered in our data.

Besides allowing for the possible heterogeneity across firms one could also relax the assumption of the constancy of the slope parameters in the time dimension. Then we would be discussing the issue of "stability"

rather than "heterogeneity." It turns out that "instability" may be the main problem with our data rather than heterogeneity and we shall explore this a bit towards the end of this paper. The original emphasis of most of our work in this paper, is, however, on the appropriate characterization of the "individuality" of our firms.

In the next section we shall outline briefly the various specifications of coefficient heterogeneity considered by us and the interrelation between them. In section 3 we shall take a first look at our major results: the observed variances of the individual parameters and estimates of their "true" heterogeneity. In sections 4 and 5, we shall estimate these variances directly within the framework of the random coefficients model, using maximum likelihood methods and the more flexible approach of regressing the squares and the cross-products of residuals on comparable squares and cross-products of the independent variables. In the last section, we look for other ways of testing the "reality" and relevance of our findings, and try to provide interim conclusions.

2. THE GENERAL FRAMEWORK

We estimate a number of different specifications of coefficient heterogeneity using both level and first differences data. Our general framework can be described by the following simple equation:

$$y_{it} = \alpha_i + \beta_i x_{it} + \epsilon_{it}$$

where y_{it} is logarithm of output per employee of firm i in year t , x_{it} is the logarithm of capital per employee, and ϵ_{it} is a disturbance distributed independently of the x_{it} 's. In this formulation the intercept parameters

α_i and the slope parameters β_i are allowed to differ across firms and different versions of the model correspond to different assumptions as to how they differ. There is also a parallel first differences formulation:

$$y_{it} - y_{it-1} = \beta_i(x_{it} - x_{it-1}) + \epsilon_{it} - \epsilon_{it-1}$$

The different models we look at can be summarized as follows:

<u>Levels</u>	<u>First Differences</u>
A. $\alpha_i = \alpha, \beta_i = \beta$	a. $\beta_i = \beta$
B. $E\alpha_i = \alpha, E(\alpha_i - \alpha)^2 = \sigma_\alpha^2, \beta_i = \beta$	b. $E\beta_i = \beta, E(\beta_i - \beta)^2 = \sigma_\beta^2$
C. $E\alpha_i = \alpha, E\beta_i = \beta; \sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha\beta}$	c. β_i distinct and fixed.
D. α_i distinct and fixed, $\beta_i = \beta$	
E. α_i and β_i distinct and fixed.	

These different combinations of assumptions can be described as follows: (A) both coefficients constant across firms (the usual linear regression assumption on pooled data); (B) random intercepts and constant slopes (the simple error components or random effects model); (C) both intercepts and slopes random across firms (the random coefficients model); (D) firm specific intercepts and a common slope parameter (the fixed effects or covariance model); (E) firm specific, fixed and different, intercepts and slopes (the fixed coefficients model). The content of the distinction between "random" and "fixed" or firm "specific" effects is partly in the assumption that the former are independent of the other variables in the equation, the x's (i.e., the firm effects are "uncorrelated," $E(\alpha_i|x) = \alpha$);

an assumption that is not imposed on the data in fitting the fixed effects versions of this model. Note that the uncorrelation assumption applies to the absence of a relationship between the individual parameters and the individual x 's and not to the absence of a relationship between the parameters themselves.⁴ We will indeed be estimating a covariance between the intercept and slope parameters also in the "random" effects model but our main emphasis will be on the slope parameter β and its variance.

Since differencing eliminates the individual intercepts, we are left with only three comparable first difference versions, all with no intercepts: (a) a single common slope; (b) randomly differing slopes; and (c) firm specific slopes. Versions E and c, which allow for complete heterogeneity across firms, are the most general ones and all the other versions are nested within them and could be tested for as different restrictions on these more general models. Note that we have been particularly explicit about the structure of the variance-covariance matrix of the disturbances (Σ). In most of what follows we shall assume that there is no serial correlation (or heteroskedasticity) beyond that which arises from the presence of the individual firm effects which are to be modeled explicitly; or that the serial correlation has been adequately eliminated or reduced by the first difference transformation. The issue of the correct specification of Σ is considered in the Appendix where we try to deduce the correct sampling errors for our individual firm estimates. Note also that we have not allowed for time variation in these parameters. The variability over time in the α_i would be absorbed in the ϵ_i . Allowing β_i also to vary over time would strain our already overextended

computational framework.⁵ But we shall come back to this possibility in discussing some of our results below.

3. A FIRST LOOK AT THE MAJOR RESULTS

Table 2 presents the results of estimating the simple production function separately for each firm. It lists the averages and observed dispersions of the estimated capital coefficients across firms both for the "level" version, which is now a "within" equation with a separate constant and slope for each firm, and the first difference version which allows, to some extent, for serial correlation in the residuals from these equations. For comparison purposes we list also similar statistics for five other variables [which are not regression coefficients]: the annual levels [in logarithms] and the annual growth rates of labor productivity and of capital intensity, and the "share of capital" in value added which could be thought of as another independent estimate of the capital coefficient. The first row of the table thus gives the averages for the individual firm capital coefficients and for the firm means of the various variables; the second, their observed dispersion in the data. The average values for the estimated coefficients are remarkably close to their least squares counterparts in Table 1, the "within" levels and the "total" first differences ones, but their dispersion is immense. The estimated standard deviations do not provide, however, an adequate impression of this dispersion. At this point a picture is worth more than several paragraphs of description.

Figure 1 plots the distribution of the capital coefficients for France [the pictures are very similar for Japan and the U.S.] against their

estimated precision [the inverse of their standard errors]. The estimated range is enormous: for a coefficient with a prior mean of about 0.3 and a reasonable range between .1 and .5 we get an actual range from -1.7 to +2.3, with more than a third of the estimates being actually negative!

We did expect to see some heterogeneity at the individual level and did assume that it would be magnified by the large sampling errors arising from the rather short time series available for their estimation, but we were not really prepared for this amount of dispersion. We turned then to an attempt to determine more precisely the amount of "pure" sampling variance in the results and the complementary effort to determine the "real" amount of dispersion in these estimates and an exploration of the sense in which this dispersion is indeed "real," if any.

The next two lines in Table 2 attempt to provide a first round answer to this question. Line 3 lists the estimated "true" dispersion or heterogeneity of the β_i coefficients based on the methodology described in Swamy (1970 and 1971). Among the different estimates of coefficient heterogeneity, the Swamy estimates are the most straightforward to obtain. They are computed as the difference between the observed variance of the individually estimated firm coefficients and an average of the corresponding sampling variances. Being based on the individual regression estimates, they are robust to the possibility of correlated effects (between the α_i and β_i and the x's) contrary to the maximum likelihood estimates and those based on the regression of the squares and cross-products of the residuals. But since these estimates are computed as a difference, they may yield negative values, and may be seriously affected by even small biases in the estimates of the observed and sampling

variances. This would be the case, in particular, if the errors are heteroskedictic and/or serially correlated. A positive serial correlation in both the errors and the x's can account for part of the discrepancy that we see between our Swamy estimates of σ_β in levels and in first differences. (See the Appendix for a discussion of the sensitivity of the Swamy estimates to the misspecification of the variance-covariance matrix of the errors.)

However, what is most striking in our Swamy estimates is their order of magnitude: σ_β of about 0.20 to 0.25 in first differences ($\sigma_\beta^2 \sim 0.05$) and between 0.30 to 0.55 in levels (σ_β^2 from 0.10 to 0.30). These estimates are still much too large to be indicators of a credible amount of heterogeneity given our prior about the possible magnitude of the capital elasticities ($\beta \sim .3$, $\sigma_\beta \sim 0.1$). It is interesting to stress, in contrast, the much smaller estimated "true" dispersion in the capital share: σ_β of only 0.10 ($\sigma_\beta^2 \sim .01$) and the effectively zero estimate for the "true" dispersion in the growth rates of labor productivity and capital intensity. As far as these two last variables are concerned, Gibrat's Law applies and we find not race of persistent individual firm differences in them: all the observed variability is either common to all firms within an industry or is transitory.

The fourth line in Table 2 lists the F statistics for the hypothesis of equality of these coefficients [variables] across firms. Given the large number of degrees of freedom in our data all these F statistics are very "significant" at conventional significance levels [the critical F ratios barely exceed one for our sample sizes]. One can question, however, whether they should be used in this fashion in such large samples, reducing

the probability of Type II error as our sample sizes increase but keeping the Type I error constant. A more symmetric treatment of the inference problem, advocated by Leamer (1978), would lead to the use of much higher values for the "critical" F (about 11 for our data) and much less certainty in rejecting the hypothesis of homogeneity.

Another way of getting an impression of the "significance" of such numbers is to compare them to the F's computed for the between firms dispersion in labor productivity, capital intensity, and capital share. These are on the order of 90 and 11 to 21, respectively, as against 3 to 4 for the capital elasticity. Now that is heterogeneity!

A related view of the testing problem is indicated by Figure 2 which plots the frequency distribution of the estimated t-ratios for the individual capital coefficients and superimposes on it the theoretically expected t-distribution for this sample size (11 degrees of freedom). It is clear from this picture that the rejection of the homogeneity assumption comes from the presence of too many extreme values and is not the result of just a few outliers. We can use it also to ask the heterogeneity question in a slightly different multiple comparisons framework (see Scheffe (1959) for its relation to the standard F test): since we are looking at a large number independent test statistics, what is the right critical value for a single statistic given that we want to keep a combined 5 percent significance level for the testing process as a whole? For (approximately) 500 comparisons, as in the French sample, the individual significance level should be set at about .0001 or a critical t-value of about 6.3 instead of the conventional 2.2 level. A single observation of a t-ratio in excess of 6.3 would be enough to reject the hypothesis that all the observations

arise from the population with the same mean and differ only from each other because of sampling variability. But in the French data we have 12 such observations!⁶

This first round look at the dispersion in our data leaves us with rather ambiguous conclusions. The observed dispersion in the estimates of individual firm capital coefficients is very large and not really credible as an indicator of "reasonable" heterogeneity. The idea that the results could be due just to sampling variability is not supported by standard statistical tests, but the level at which the observed statistics fail the homogeneity hypothesis is not particularly impressive. The latter point of view is reinforced by the fact that the estimates of sampling error and the associated F-statistics, as we already noted, are very sensitive to the underlying distributional assumptions and especially to the assumption of homoskedasticity and serial independence (see Appendix).

4. MAXIMUM LIKELIHOOD ESTIMATES

In Tables 3 and 4 we present several ML estimates for our three samples and our different specifications. We estimated the specifications B, C, and b, using the maximum likelihood framework. In each case the log likelihood function is the sum of individual firm likelihood values expressed in terms of the summary statistics for each firm and the parameters to be estimated: α , β , σ_{α}^2 , σ_{β}^2 , $\sigma_{\alpha\beta}$, and σ_{ϵ}^2 (see Swamy 1971, sect. 4.3.d). For the constant parameters and the fixed effects cases, the maximum likelihood estimates are equivalent to different OLS versions and we have used the latter. For D they are the "within" estimators and for the E and c specifications they are the previously described individual

firm estimates. These ML estimates, contrary to those presented in Table 2, assume either the independence of the β_i 's or the independence of both α_i and β_i from the x 's, i.e., the absence of "correlated effects." They also assume that ϵ and the coefficients α and β are distributed normally, or more generally that their distribution belongs to the linear exponential family (see White 1981, 1982; Gourieroux, Monfort and Trognon 1984). The standard errors of the estimates under the normality assumption (computed from the inverse of the information matrix) are given in the first line of parentheses; the more general standard errors robust under heteroskedasticity and non-normality are given in the second line of parentheses. Column 2 gives the estimated values of the log likelihood, allowing one to compute directly the various likelihood ratio (LR) tests for the randomness of the parameters (i.e., the nullity of the variances σ_β^2 and/or σ_α^2) or for their equality across firms (i.e., the equality of β_i and/or α_i for all i). We shall turn first to the tests of heterogeneity, before looking at the orders of magnitude themselves. Table 5 recapitulates the 2LR values per degrees of freedom p of the tests (where p is the number of restrictions or additional parameters) and the corresponding critical values of the χ^2 or F statistics at the one percent level of significance: $\chi^2_{0.01}(p)/p = F_{0.01}(p, \infty)$.

The pattern of the results is quite similar for all three samples. All of the rejections appear to be very significant statistically. There are, however, large differences between them, if one takes an eclectic view of comparing gains in likelihood per additional parameters. In this sense, the largest benefit comes from relaxing the constancy of the intercept parameter α . With only one additional parameter σ_α^2 , the random effects

specification is extremely parsimonious, this one parameter being "equivalent" to about 80% of the (N-1) additional parameters of the fixed effects specification. The additional improvement arising out of the relaxation of the constancy of the slope parameter, although less dramatic, is still quite sizeable. Again the random coefficient specification performs relatively well (the additional variance parameter σ_{β}^2 is "equivalent" to about 30 to 40 percent of the residual variance reduction achieved by the N-1 firm specific coefficients β_i in levels, and about 15 to 25 percent of it in first differences).

The LR tests of equality of the slope parameters β (D if E and a if c are equivalent to the usual F tests (Chow tests of stability or homogeneity) that have been discussed in the introduction (with our large number of observations, 2LR/p values are in fact equal or almost so to F values; they differ here only because the F given in Table 2 is based on a different weighting). As we have noted already, although they are very "significant" at conventional significance levels, they are not all that "high" if one takes into consideration our large sample sizes and follows Leamer's view of using a critical value which increases with the number of degrees of freedom.

We are, however, primarily interested in the order of magnitude of our heterogeneity estimates. The more relevant numbers are the estimated variances (σ_{β}^2) for the two random coefficient specifications C and b. The figures in brackets in Tables 3 and 4 given for the fixed coefficient specifications are the empirical variances of the firm-specific estimated β_i . As expected they are much greater than the estimated "true" σ_{β}^2 , most of the differences corresponding to the sampling variances (i.e., the

variances due to the errors in the model) and our limited study period. Although unaffected by the imprecision of the individual estimation of the β_i , the estimated σ_β^2 remains large, and far too large for numbers which a priori we would expect to be of the order 0.01 ($\sigma_\beta = 0.10$) and not to exceed 0.04 ($\sigma_\beta = 0.20$).

There is a rather clear pattern in this regard in the results for the three samples. The σ_β^2 's estimated in first differences are much smaller than those estimated in levels, but they still seem too large, being about 0.06 in France and Japan, and 0.04 in the U.S. (the corresponding estimates in levels are about 0.15, 0.21 and 0.07, respectively). The reason why the variances are so different when estimated in levels and in first differences (and nearer to what we would deem acceptable for the latter) is unclear. A plausible explanation is, of course, the possibility of correlated effects, since maximum likelihood estimation in levels assumes that the random intercept effects α_i are uncorrelated with the capital-labor ratio. Going to first differences disposes of that problem. The fact that our estimates of the (mean) capital elasticity β tend to be higher (clearly for Japan, more or less so also for France, but rather the opposite for the U.S.) in specifications which are affected by the presence of correlated effects (A, B, and C) is another indication of the same phenomenon. There may be also some coherence in the observed discrepancies in the estimated β and σ_β^2 in levels and first difference: they are both quite wide for Japan and both modest for the U.S., France being in a somewhat intermediate position.

Another explanation for the discrepancies in the σ_β^2 (not the β) is the positive correlation of the errors ϵ in the regression, and hence their

negative autocorrelation in first differences. This would cause a positive bias in the level estimates of σ_{β}^2 , since the x themselves are strongly positively correlated, but only a negligible one in the first difference estimates, the differenced Δx 's being only weakly serially correlated.

5. ESTIMATES FROM REGRESSIONS ON SQUARES AND CROSS-PRODUCTS OF RESIDUALS

There is another way of estimating the variance components of heterogeneity suggested by the random coefficients model literature. This method (developed among others by Hildreth and Houck (1968), Goldfeld and Quandt (1972), Amemiya (1977) and more recently by MaCurdy (1981, 1985)) relies on the fact that the variances and covariances of the dependent variable y , conditional on the explanatory variable x , are linear functions of the relevant squares and cross-products of the x 's. Thus, if the mean parameters of the model have been estimated consistently in a first stage and the corresponding residuals u_{it} computed in this stage, it can be shown that regressing their squares and cross-products on the corresponding squares and cross-products of the x 's yields consistent estimates of the variance parameters. Since these different regressions constitute a system of related equations (with appropriate cross-equations restrictions on the parameters), they can be estimated more efficiently (given the mean parameters) by quasi generalized least squares (QGLS).

Letting u_{it} and Δu_{it} be the residuals from the simple OLS estimates of our model in levels and in first differences (A and a), we have two systems of equations across firms (the observations being over i). The first one in levels consists of T equations on squares, of $(T-1)$ on cross-products lagged by one year, of $(T-2)$ on cross-products lagged by two years, ... or a

total of $T(T+1)/2 = 91$ of them. Similarly in first differences there are such $(T-1)$ equations on squares, ..., for a total of $T(T-1)/2 = 78$. Assuming that ϵ 's are uncorrelated over time, they can be written respectively as:

$$u_{it}^2 = \sigma_{\beta}^2 x_{it}^2 + 2\sigma_{\alpha\beta} x_{it} + (\sigma_{\alpha}^2 + \sigma_{\epsilon}^2) + v_{it}$$

(SCP)

($t=1, \dots, T$)

$$u_{it} u_{it'} = \sigma_{\beta}^2 x_{it} x_{it'} + \sigma_{\alpha\beta} (x_{it} + x_{it'}) + \sigma_{\alpha}^2 + v_{itt'}$$

($t \neq t'$; $t=1, \dots, T$; $t'=1, \dots, T$)

$$\Delta u_{it}^2 = \sigma_{\beta}^2 \Delta x_{it}^2 + \sigma_{\Delta\epsilon}^2 + w_{it}$$

(SCP')

($t=1, \dots, (T-1)$)

$$\Delta u_{it} \Delta u_{it'} = \sigma_{\beta}^2 \Delta x_{it} \Delta x_{it'} + w_{itt'}$$

($t \neq t'$; $t=1, \dots, (T-1)$; $t'=1, \dots, (T-1)$)

We have estimated these two systems by OLS and QGLS (but shall report here only on the latter). Although somewhat cumbersome (because of the number of equations), this approach is more flexible than the ML method. It is easier to investigate various assumptions about the error terms or the random parameters themselves within this framework. In particular, we can accommodate serial correlation of the errors by not imposing the

assumption that the intercepts of the different cross-product equations of the system in levels are equal and allow also for non-stationarity of the errors by not assuming them to be equal within various groups of equations. Similarly, we can allow for unsynchronized random shocks on our response parameter β by relaxing the equality constraints on the coefficients of the squares and cross-products of the x's in different ways.

It is also possible to use first stage residuals corresponding to alternative variants of the (first order) specification of our model. In the case of the French sample, we have allowed for the possibility of systematic (synchronized) changes over time in the capital elasticity β , and for correlated (with the x's) firm effects by including all the available x_{it} in the first stage regressions (using what is known as Chamberlain's Π matrix approach, 1984).

The results of our main computations are summarized in Tables 6, 7, and 8, respectively, for the fully constrained estimates, and for those constrained only within groups of equations and for those obtained from the different year squares equation and from four years apart cross-products equation. Compared to our previous ML estimates, there is a pleasant surprise: the fully constrained estimates of σ_{β}^2 in levels are much smaller, and now they are also quite close to the first difference estimates. Both types of estimates, however, stay on the high side of what would seem a priori reasonable: the estimated σ_{β}^2 range from 0.03 ($\sigma_{\beta} = 0.18$) for the U.S. in first difference, to 0.06 ($\sigma_{\beta}^2 = 0.25$) for Japan in levels.

While these results do not change our main conclusions significantly they do raise some new puzzles. It is not clear why our earlier ML

estimates of σ_{β}^2 in levels and those summarized in Table 6 should be so different: in particular they are both affected similarly by correlated intercept effects and by serial correlation of the errors (at least if we impose the full set of equality constraints on the intercepts in the system of square and cross-product regressions). It is also puzzling why our new estimates of σ_{α}^2 (in levels) and of σ_{ϵ}^2 (in levels and first differences) are smaller than their ML counterparts. These anomalies point to some, still unclear, misspecifications in our model.

The incompletely constrained estimates (in Tables 7 and 8) may shed some light on this mystery. They show that there is much instability in our results, especially in first differences. The example of the French sample is most striking. We find a first difference estimate of σ_{β}^2 of about 0.12 for the squares equations, with corresponding estimates for individual year equations varying from 0.08 up to 0.42 (in 1974/73). This estimate goes down to 0.04 in the four year apart cross-products equations with corresponding separate estimates ranging again from -0.18(!) to 0.16. It drops down again to -0.04 (!) in six year apart cross-product equations, and up again to 0.10 in the eleven year apart cross-product equations. For the other two samples, the pattern is roughly similar, though less extreme. The estimates of σ_{β}^2 in levels are much more stable; they do indicate, however, a clear downward trend when we go from those estimated on squares to the estimates based on the twelve year lagged cross-products. This could be due to some form of random shocks in β_{it} , unsynchronized but strongly correlated over time, in addition to the fixed heterogeneity component β_i . Such shocks unfortunately do not seem to account for what we

see in first differences, and we have not been able (yet) to characterize the instability in these estimates satisfactorily.

6. SWAMY ESTIMATES

Among the different types of estimates of coefficient heterogeneity, the Swamy estimates are the most straightforward to obtain. As already noted, they are computed as the difference between the observed variance of the individually estimated firm coefficients and an average of the corresponding sampling variances. Since the Swamy estimates are based on the individual estimates, they are robust to correlated effects. However, being computed as a difference, they may be affected by even small biases in the estimated sampling variances, which could be the case if the errors ϵ in our model are heteroskedastic and/or serially correlated. In particular, the serial correlation both in the errors and the x 's will account for part of the discrepancy between the Swamy and ML estimates of σ_{β}^2 in levels and those in first differences. In order to have a clearer appreciation of the sensitivity of the estimates to these (second order) misspecification, we have computed the Swamy estimates under alternative assumptions for the variance-covariance matrix of the errors.

7. EXTENSIONS AND INTERIM CONCLUSIONS

It is not easy to summarize our various results. There are two numbers worth keeping in mind: our prior expectation about a reasonable heterogeneity in the true β 's is a standard deviation of 0.1. The observed standard deviation at the individual level is about 0.5 (0.4 in first differences). A direct subtraction of an estimate of the sampling variance

(the Swamy method) results in a "residual" estimate of the true dispersion σ_β between 0.4 in levels and 0.2 in first differences, the latter still being about twice as high as our prior expectations. More "direct" estimates of σ_β which associate it with the coefficient of the squared x's (or their crossproducts) in the squares and crossproducts of residuals regressions or their maximum likelihood analogues, however also, at their lower range, around 0.2. Thus, allowing for sampling variability or estimating the dispersion directly, still yields estimates that are "too high."⁷

At this point there are two possibilities: we could accept the 0.2 estimate as real and revise our expectations. If we allow for the fact that our simple production function equation may be seriously misspecified (there may be decreasing or increasing returns to scale, capacity utilization fluctuations may matter, capital may be mismeasured, etc.), then it may be reasonable to expect the β in a misspecified equation to have a different prior distribution. For example, a more agnostic view would still keep it between zero and one but would not be surprised by a standard deviation of 0.2. An alternative, though not mutually exclusive view, would lead us to inquire into the "reality" of this heterogeneity and its stability over time. Unless the misspecification is in some sense permanent, it is not particularly interesting.

We have actually explored a number of first order extensions of this model and found that they do not reduce the estimated σ_β significantly. Neither letting go of the constant returns assumption or using value added output measures and allowing thereby for the presence of the materials input (in French and Japanese data) changes our results significantly. The

"improvement" in the specification is counterbalanced by the decline in the "net" variance of the capital-labor ratio from which the β 's have to be estimated.⁸ The impact of capacity utilization fluctuations remains still to be explored because we do not have a good handle on it in our data. Moreover, its effect is unlikely to be "stable," it would differ from period to period, which leads us to our second way of looking at this problem.

We also worried about the possibility that the observed heterogeneity is just an artifact of a few outliers that using some alternative more "robust" approach may make it go away. A look at Figure 1 will convince one that our results do not depend just on a few outliers. Nevertheless, we pursued two approaches towards robustness. In one approach, we eliminate all firms whose within variance in X was either very small or extremely large. This had essentially zero effect on our estimated "true" σ_{β}^2 , it reduced the observed variance and the estimated sampling variance about equally. The other approach, computing a robust estimate of the observed variance of the β 's from their interquartile range and subtracting the median of the estimated sampling variances rather than their mean resulted also in only minor changes in the final estimates, reducing somewhat the σ_{β}^2 estimated in levels, from .19 to .15, but raising it for the first differences based estimates, from .05 to .08. (The reported numbers are for Japan but the results are similar in all three countries.) It is clear that the "over-dispersion" that we observe (see Cox 1983 for related use of this term) is not the result of a few errant firms, but is actually a pervasive aspect of our data and our approach to them.

Why do we care about heterogeneity? If people differ in their response coefficients and if we contemplate some experiment which will impinge differentially on them, then it would be valuable to know how they differ and how this is related to any policy change we may wish to pursue. If they do differ but these differences are not permanent, then knowing their heterogeneity today may not help us any tomorrow. This leads us to question also the time stability of the β 's and not just their cross-sectional heterogeneity. One way of asking this question is to divide our period into two, redo the analysis, and investigate the question whether the estimated heterogeneity is the "same" in both periods, whether the estimated within period β 's are correlated with each other.⁹

The results of reestimating our model for each firm in each of the countries, separately for the two subperiods: 1967-1973 and 1974-1979, and computing the correlation between the estimated $\hat{\beta}_1$'s in the two periods are given in Table 9. These correlations (the last two columns of row 4) are remarkably low, on the order of .05 to .1, implying a rather substantial time instability in these estimates and throwing serious doubt on the "reality" of the previously estimated heterogeneity levels. Row 3 shows the associated estimates of σ_β based on their between periods covariance. They are much lower than the total period based estimates and are now much closer to what we might think is a priori reasonable. Moreover, the levels and first differences based estimates are now very close to each other. For comparison purposes we also list similar numbers for the average output-per-man ratios and capital shares. The heterogeneity in them is much more "permanent" and so is also the heterogeneity in the intercepts (not shown here). For example, in France the between periods "permanent"

variance accounts for 95 percent of the total for the average output-per-man ratio and for 90 percent of the total variance in the average capital share. At the same time, the estimated variance of the β 's from the between periods covariance is only 6 percent of the total period estimates in levels and 28 percent in the first differences based calculations. It is obviously a much less stable aspect of average firm behavior.¹⁰

The most common way of allowing for heterogeneity in such data is to estimate separate equations for different industries. We have already done something along these lines by allowing separate industry-time constants. But we can ask the additional question: Does the estimated heterogeneity have any "industrial" content? A priori, we would think that it should have some. Otherwise, the "reality" of the estimated dispersion in the β_i 's may be doubtful and also the usefulness of knowing it, even if it were present. To check on this, we can run an analysis of variance of the estimated individual slope coefficients using the same 14 industries breakdown which we have already used in defining the time-industry dummy variables. We have computed this for our sample of Japanese firms with the result that the industrial structure, in the form of 14 industry dummy variables, accounts for less than 5 percent of the observed variance in the $\hat{\beta}_i$'s in first differences, and even less in levels (less than 2 percent). Hence, consistent differences between industries in the parameters of the production function are not significant contributors to the firm heterogeneity as estimated by us.

Another way of asking about the "reality" of the observed dispersion in the β 's is to ask how they are correlated with other estimates of the

same thing. Under the assumption of competition and constant returns to scale the capital share should roughly equal β . Even if one relaxes these assumptions and allows for deviations from purely competitive conditions or constant returns to scale, one would still expect a significant positive correlation between these two measures. The facts summarized in line 5 of Table 9 are, however, rather different. The observed correlations are insignificant. Even if one takes the maximal number of .04 for France and adjusts it upward for the attenuation due to sampling error in these averages, one would still get only .05 for this correlation, which is much too small to allow us to think of one of these measures as being a reflection of the other. It appears that the estimated dispersion in the β 's has little to do with what we may think are the reasonable sources for it.

Having been alerted to the possibility of time-instability in these estimates we can now also look for it in the regression based estimates of their variances. Unless the β 's are stable over time, the regressions based on the crossproducts of the residuals across different years and the crossproducts of the x's, rather than on their squares, should yield much smaller estimates. This would be true because the time instability in the β 's should show up in the residual but not be strongly correlated over time. Since this effect may be masked by the presence of correlated effects in the data (the correlation of the intercepts with the x's, which is assumed away in such estimation), only the first differences or II matrix residuals based calculations are relevant here.¹¹ Looking at the last three columns of Table 7 indicates clearly that the crossproduct based variance estimates are much lower and unstable. The parallel II matrix

residual based results summarized in Table 10 also indicate that when one estimates this variance from "distant" cross products one gets a much lower estimate of it. In short, whatever it is that these individual β 's measure it is not a very stable aspect of firm behavior.

Moreover, there is some evidence that this instability is year specific. If we break our data into even and odd years, rather than the early and later periods, the correlations between the β 's are much higher: .5 versus .05 for the two periods cut (in France). Thus, the fluctuations in β are not random from year to year. This can also be seen in Table 8, where we exhibit the separate estimates of σ_{β} based on squared residuals from different years. It is quite clear that something happened around 1974 which is not very well captured by the model or by allowing for heterogeneity in the β 's across firms rather than across time.

What can we conclude about the average of our estimated β 's, the capital elasticity of output, rather than its dispersion? The major conclusion is that we have not solved the between-within puzzle. Heterogeneity is not the answer to it. This can be seen by comparing the means of the individual β 's to either the between or the within estimates which are reproduced, for convenience sake, in Table 12. Except for the U.S., lines 1 and 2 differ significantly (as do also lines B and D in Table 3), but lines 2 and 3 do not. Line 3 is the unweighted average of the individual β 's, line 2 is the weighted average and the two are remarkably close, implying that the weighting scheme of the "within" estimator is not the culprit for its "lowness." There do appear to be real correlated effects (between the α_i , the individual constants, and the x 's) in the cross-sectional dimension (at least in France and Japan). There is

no evidence, however, of a significant correlated effect in the β 's. Line 3 (or 5) which does not assume that the β 's are uncorrelated with the X's turns out to be essentially the same as line 4 (or 6) which imposes the no-correlation assumption on these estimates. Line 3 is also effectively the same, as we have already seen, as line 2, which weights these estimates by the individual firm variance of the x's.¹²

In a sense this evidence of no real relation between the individually estimated β 's and the capital-labor ratios is another indication of the "unreality" or irrelevance of their estimated dispersion. What is impressive in this table is, however, the closeness of the various estimates of the mean β among the set of estimators which allow for a dispersion in the α 's. Even noting the slight reversal for Japan and U.S. in the lower part of this table which recapitulates some of the first-differences based results, there is a sense in which the average response coefficients appear to be rather well estimated.

What is then the answer to the rhetorical question in the title of this paper? It is negative given the current status of our data and the variables available to us. Without data on factor prices at the micro level and information on actual capacity utilization, it is unlikely that we could derive stable production relationships at the individual firm level. Moreover, the simple production function model, even when augmented by additional variables and further non-linear terms, is at best just an approximation to a much more complex and changing reality at the firm, product, and factory floor level. While it is possible to get reasonable estimates of the average response coefficient which are consistent across different data sets and different estimation methods, there does not appear

to be much promising content in the estimated dispersion. In this sense, if we are interested in the average behavior of productivity then "aggregation", or more correctly pooling, i.e., estimating a common coefficient across a large number of firms, may not be "necessarily bad" (cf. Grunfeld and Griliches, 1961).

Much more could be said about the various estimates than we have been able to do here. We could do more exploring of the differences that arise from the assumption of correlated individual effects ("random" versus "fixed" effects estimates), from different treatments of serial correlation in the disturbances, and from the results of alternative estimation methods (see the Appendix for some additional discussion of these topics). But the main problem facing us is the inadequacy of the underlying simple production function model. We need to figure out ways of allowing for the discrepancy between recorded inputs and actually used levels, for lags, and for the simultaneity in the various behavior relations associated with the production process. As always, much work remains to be done, but then one needs also to start somewhere.

Footnotes

*ENSAE and EHESS, Paris and Harvard University and NBER, respectively. We are indebted to the National Science Foundation (SES-82-0806), CNRS, the Sloan Foundation, and the NBER Productivity Program for financial support. We are grateful to Thomas Abbott, Jean Marie Chanut, and Ghazala Mansuri for research assistance and to Bronwyn Hall for many useful comments.

1. For examples of this work see Griliches and Mairesse, 1983, 1984, and 1985, and Mairesse, 1987.
2. What is visible in this paper is only the tip of an iceberg of a great deal of work with these data. There are many interesting but tangential issues, both substantive and econometric, which could be taken up as we traverse this range of topics.
3. Using the notations and framework given in the next section, the between estimates can be written as

$$\hat{b}_b = \sum_i b_i d_i^2, \text{ where}$$

$$\hat{b}_i = (\bar{y}_{i.} - \bar{y}_{..}) / (\bar{x}_{i.} - \bar{x}_{..}) \text{ and } d_i^2 = (\bar{x}_{i.} - \bar{x}_{..})^2 / \sum_i (\bar{x}_{i.} - \bar{x}_{..})^2$$

While the within estimates are equivalent to

$$b_w = \sum_i \hat{\beta}_i w_i^2, \text{ where}$$

$$\hat{\beta}_i = \sum_t (x_{it} - \bar{x}_{i.})(y_{it} - \bar{y}_{i.}) / \sum_t (x_{it} - \bar{x}_{i.})^2$$

$$\text{and } w_i^2 = \frac{\sum_t (x_{it} - \bar{x}_{i\cdot})^2}{\sum_j \sum_t (x_{it} - x_{j\cdot})^2}$$

The argument in the text assumes that $E\hat{b}_i = E\hat{\beta}_i = \beta_i$, which implies the absence of "correlated effects" ($\hat{\alpha}_i$'s related to the x 's). $E(b_w)$ will be less than $E(\hat{b}_b)$, if there is a negative correlation between the β_i 's and the w 's.

4. Since the x 's are also economic choice variables, this assumption may not be reasonable for either the α 's or the β 's. In a non-constant returns case, with x (the capital-labor ratio) fully variable, it would be proportional to $\beta_i/(1 - \beta_i)$. This could introduce a correlated effects problem in the cross-sectional estimates, but would not be a problem in the time-series dimension, where most of the variability in x would come from changes in the relevant factor prices.

5. See Hsiao (1986) for more elaborate versions which allow for variability in both dimensions.

6. Both figures 1 and 2 plot the levels based β_i . The corresponding graphs for the first differences based β_i are similar in shape but the dispersion is smaller. In the French sample, there are only three β_i 's whose t -ratios fall outside the 6.3 range.

7. It is possible to speculate at length why different methods of estimation give us different results. We have done so a bit already in discussing them. It is basically an indication of some additional misspecification in our model, a point to which we shall come back below.

8. This can be illustrated by considering parallel results which do not include separate time-industry constants, putting all of the trend effects into the capital-labor ratio. Such estimates yield, obviously, a much

higher $\bar{\beta}$ (e.g., 0.41 versus 0.16 in first differences for Japan), but also a much smaller dispersion (an observed σ_{β} of 0.31 versus 0.43) and the implication of no true heterogeneity, all of the observed dispersion being attributable to sampling variability. Thus, adding statistically important variables, the sectoral year-dummies, and reducing thereby the potential misspecification of the model, increases the estimated heterogeneity of β 's instead of reducing it.

9. One could, of course, pursue more general models of true instability, letting $\beta_{it} = \bar{\beta} + \delta_i + \lambda_t$ or $\beta_{it} = \bar{\beta} + \delta_i \lambda_t$. See Hsiao (1986) and Cragg (1987) for such extensions. There are also more general approaches possible to the definition and testing of parameter heterogeneity and stability. See Dutta and Leon (1986) for example.

10. The correlation between the constants estimated separately for the two periods is 0.45 for France. Adjusted for their respective sampling variances, the "deattenuated" between periods correlation is 0.88, implying a large measure of stability for this aspect of firm behavior. A comparable deattenuation of the between periods correlation coefficient for the slopes raises it only from 0.05 to 0.18.

11. The correlated effects model can be written as (see Chamberlain, 1984):

$$y_{it} = \alpha_i + \beta_i x_{it} + e_{it}$$

and

$$\alpha_i = \delta_t' x_i + \eta_i$$

where x_i is the vector (x_{i1}, \dots, x_{iT}) of all the x 's in the different periods for individual i . The estimated residual from the overall Because the

residual square and cross-product equations we estimate ignore the non-corresponding square and cross-product x terms, they are likely to overestimate σ_β . First differences, which eliminate the α_i 's, or the Π matrix residuals, which include the δ_t 's terms in the equation and leave only the η_i in the residual, do not suffer from this problem.

12. These numbers are either very close (or very different) and hence we did not bother to subject them to formal Hausman (1978) tests.

Appendix: Swamy Type Estimates and Their Sensitivity to
Heteroskedasticity and Serial Correlation

Since the Swamy type estimates can be very sensitive to second-order manipulation of our model, we have computed them under alternative assumptions for the variance-covariance matrix of the errors. A worthwhile extension of this work would be based on more robust estimates of the variances involved, using nonparametric methods such as the Jackknife or Bootstrap.

We use the following matrix notation, in addition to that given in section 2 of the text: Δ is the (2,2) variance-covariance matrix of the coefficients (i.e., $\delta_{11} = \sigma_{\beta}^2$, $\delta_{22} = \sigma_{\alpha}^2$, $\delta_{12} = \delta_{21} = \sigma_{\alpha\beta}$), Σ_i is the (T,T) variance-covariance matrix of the errors ϵ_{it} for firm i ($\Sigma_i = E(\epsilon_i \epsilon_i')$), X_i is the (T,2) matrix of observations on x for firm i and the constant (i.e., $X_i = (x_i, 1_T)$), and $\underline{\beta}$ stands for the (2,2) vector of mean coefficients (i.e., $\underline{\beta} = (\beta, \alpha)'$). The individual firm OLS estimates and their corresponding sampling variances are as follows:

$$\hat{\underline{\beta}}_i = (X_i' X_i)^{-1} X_i' y_i$$

$$V(\hat{\beta}_i | X_i, \beta_i) = (X_i' X_i)^{-1} X_i' \Sigma_i X_i (X_i' X_i)^{-1}$$

The observed variance-covariance matrix of the individual firm OLS estimates is

$$\hat{V}(\beta_i) = \frac{1}{N-1} \sum_{i=1}^N (\hat{\beta}_i - \bar{\beta})(\hat{\beta}_i - \bar{\beta})' \text{ with } \bar{\beta} = \frac{1}{N} \sum_{i=1}^N \hat{\beta}_i .$$

We can show that:

$$V(\hat{\beta}_i) = VE(\hat{\beta}_i | X_i, \beta_i) + EV(\hat{\beta}_i | X_i, \beta_i) = \Delta + EV(\hat{\beta}_i | X_i, \beta_i)$$

$E(V(\hat{\beta}_i | X_i, \beta_i))$ can be estimated by the average sampling variance:

$\bar{V}(\hat{\beta}_i | X_i, \beta_i) = \frac{1}{N} \sum_{i=1}^N V(\hat{\beta}_i | X_i, \beta_i)$, assuming that ϵ_{it} are uncorrelated across firms, and thus the true variance-covariance matrix of the coefficients Δ can be estimated by:

$$\hat{\Delta} = V(\beta_i) - \bar{V}(\hat{\beta}_i | X_i, \beta_i)$$

This formula generalizes that given by Swamy, where heteroskedasticity is allowed only across firms, i.e., $\Sigma_i = \sigma_{\epsilon_i}^2 I$.

Thus, different assumptions about Σ_i , yield alternative estimates of Δ . This is conditional, however, on obtaining consistent estimates of Σ in the panel data context, with N going to "infinity" for a fixed T . The most simple case is that of homoskedasticity and independence of the errors:

$\Sigma_i = \sigma_{\epsilon}^2 I$ and $V(\hat{\beta}_i) = \frac{1}{N} \sigma_{\epsilon}^2 \sum_{i=1}^N (X_i' X_i)^{-1}$. It is also possible to allow for heteroskedasticity across firms (as Swamy originally did): $\Sigma_i = \sigma_{\epsilon_i}^2 I$, the sampling variance being computed then as the arithmetic average of the individually estimated sampling variances: $\frac{1}{N} \sum_{i=1}^N \sigma_{\epsilon_i}^2 (X_i' X_i)^{-1}$. [The Swamy estimates in Table 2 are computed in this way]. We also consider the case of heteroskedasticity within firms, where Σ_i is a diagonal matrix (White, 1980) which depends on X_i .¹ In order to deal with serial correlation, we assume a given autocorrelation (T,T) matrix of the errors TCOR corresponding to an autoregression of the first order with autocorrelation parameter $\rho = .5$ in levels (and the corresponding

autocorrelation matrix in first differences). Following Solon (1984), a close examination of the first three serial autocorrelation coefficients of the residuals computed both from the fixed effects specifications in levels and in first differences, indeed suggests that a first order autoregression with a ρ of about .5 would be an adequate description of the error process. To take into account the serial correlation of the errors, we have also used the Cochran-Orcutt transformation to compute the Swamy estimates in quasi-difference form (i.e., $(y_{it} - 0.5y_{it-1})$, and $(x_{it} - 0.5x_{it-1})$).

The various estimates of the "true" variance in the coefficients, that we can compute under these different assumptions are given in Table 12. We also present in this table the corresponding QGLS estimates of the mean β . These results are only given for the French sample, since they are very similar for our other two samples.

The direction of the potential biases that we see in the estimated σ_{β}^2 is what one might have expected, but their magnitude is more of a surprise. Allowing for heteroskedasticity across firms or within firms increases the estimated σ_{β}^2 by about 0.01 to 0.03. This increase is relatively minor for the level estimates, but is more serious for the first difference estimates, since they are much smaller than the former and of about the same order of magnitude. Such bias, however, corresponds to a negative correlation (across or within firms) between the variance of the residuals and the corresponding squares of the x's (or Δx 's), which may be in part spurious (due to sampling errors or outliers: very small or very large squares of the x's or Δx 's).

Taking into account the serial autocorrelation of the errors, affects our level estimates severely, changing the estimated σ_{β}^2 from about 0.18 to 0.13. This is confirmed by the quasi-differenced estimates, which are even smaller, with estimated σ_{β}^2 of about 0.08. The first differences based estimates of the sampling variance are not, however, biased by the (negative) serial correlation in the differenced errors, because the differenced Δx 's are practically uncorrelated over time (contrary to the x 's in levels which are strongly positively correlated). It is interesting to note that there is little difference in the various QGLS estimates of β corresponding to the various weighting schemes. Contrary to the variance estimators, they are all consistent under the alternative (second order) misspecification of the variance-covariance matrix of the errors.

APPENDIX NOTES

1. There is an identification problem in discussing heteroskedasticity in the context of random coefficient models. The only heteroskedasticity that matters here is one that is a function of the x's but that, in principle, should be in the definition of β_{it} . The random coefficients model is, in fact, a particular model of heteroskedasticity. Since our main focus is on the variability of β_{it} in the i (individual) dimension, it is sufficient for us to define $\beta_{it} = \beta_i + e_{it}$, where the distribution of e_{it} may be a function of x_{it} , and throw them (e_{it}) into the definition of the overall disturbance ϵ_{it} .

References

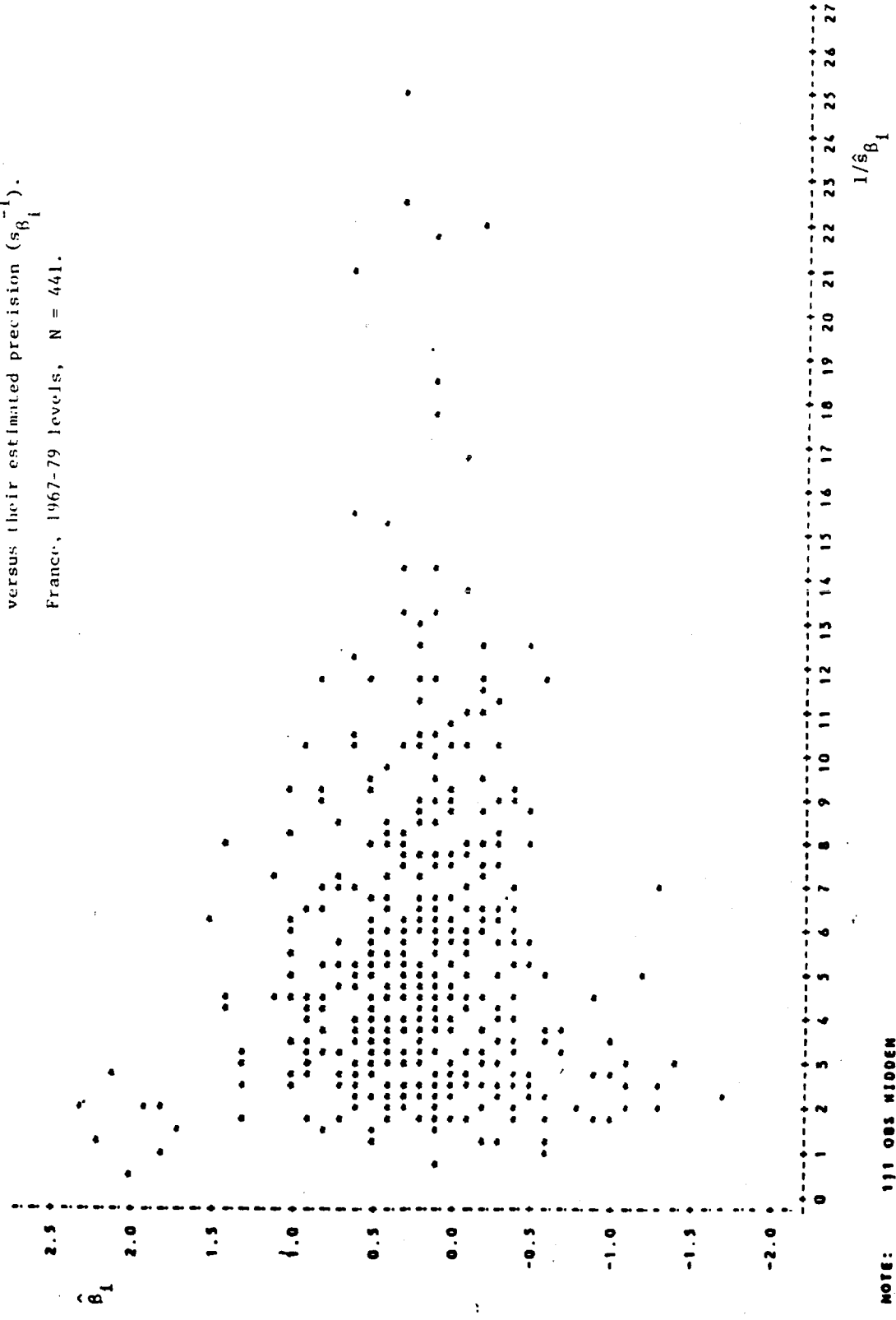
- Amemiya, T. 1977. "A Note on a Heteroskedastic Model," Journal of Econometrics 5: 295-299.
- Chamberlain, G. 1984. "Panel Data," in Handbook of Econometrics, Vol. II, ed. by Z. Griliches and M. Intriligator, 1247-1318. Amsterdam: North-Holland.
- Griliches, Z. and J. Mairesse. 1983. "Comparing Productivity Growth: An Exploration of French and U.S. Industrial and Firm Data," European Economic Review 21: 89-119.
- _____. 1984. "Productivity and R&D at the firm level," in R-D, Patents and Productivity, edited by Z. Griliches, 339-374. The University of Chicago Press.
- _____. 1985. "R&D and Productivity Growth: Comparing Japanese and U.S. Manufacturing firms," National Bureau of Economic Research Working Paper No. 1778. To appear in an "Income and Wealth" conference volume.
- Goldfeld, S.R. and R.M. Quandt. 1972. Non Linear Methods in Econometrics. Amsterdam: North-Holland.
- Grunfeld, Y. and Z. Griliches. 1960. "Is Aggregation Necessarily Bad?," Review of Economics and Statistics, :1-13.
- Hildreth, C. and J.P. Houck. 1968. "Some Estimates for a Linear Model with Random Coefficients," Journal of the American Association 63: 584-595.
- MaCurdy, T. 1981. "Multiple Time-Series Models Applied to Panel Data," National Bureau of Economic Research Working Paper No. 646.

- Mairesse, J. 1987. "Les lois de la production ne sont plus ce qu'elle
etaient: une introduction a l'econometrice des panels," Working Paper
ENSAE, to appear in La Revue Economique.
- Mundlak, Y. 1978. "On the Pooling of Time Series and Cross-Section Data,"
Econometrica 46: 69-85.
- Scheffe, H. 1959. The Analysis of Variance. New York: John Wiley.
- Swamy, P.A.V.B. 1971. Statistical Inference in Random Coefficient
Regression Models. Berlin: Springer-Verlag.
- _____. 1970. "Efficient Inference in a Random Coefficient Model,"
Econometrica 38: 311-23.
- White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix
Estimation and a Direct Test for Heteroskedasticity," Econometrica 48:
817-838.

Figure 1

Estimates of individual capital elasticities ($\hat{\beta}_1$) versus their estimated precision ($s\hat{\beta}_1^{-1}$).

France, 1967-79 levels, N = 441.



NOTE: 111 OBS WIDEN

Figure 2

Frequency distribution of the t-ratios
for the individual capital elasticities
 β_1 . France, 1967-79 levels, N = 441

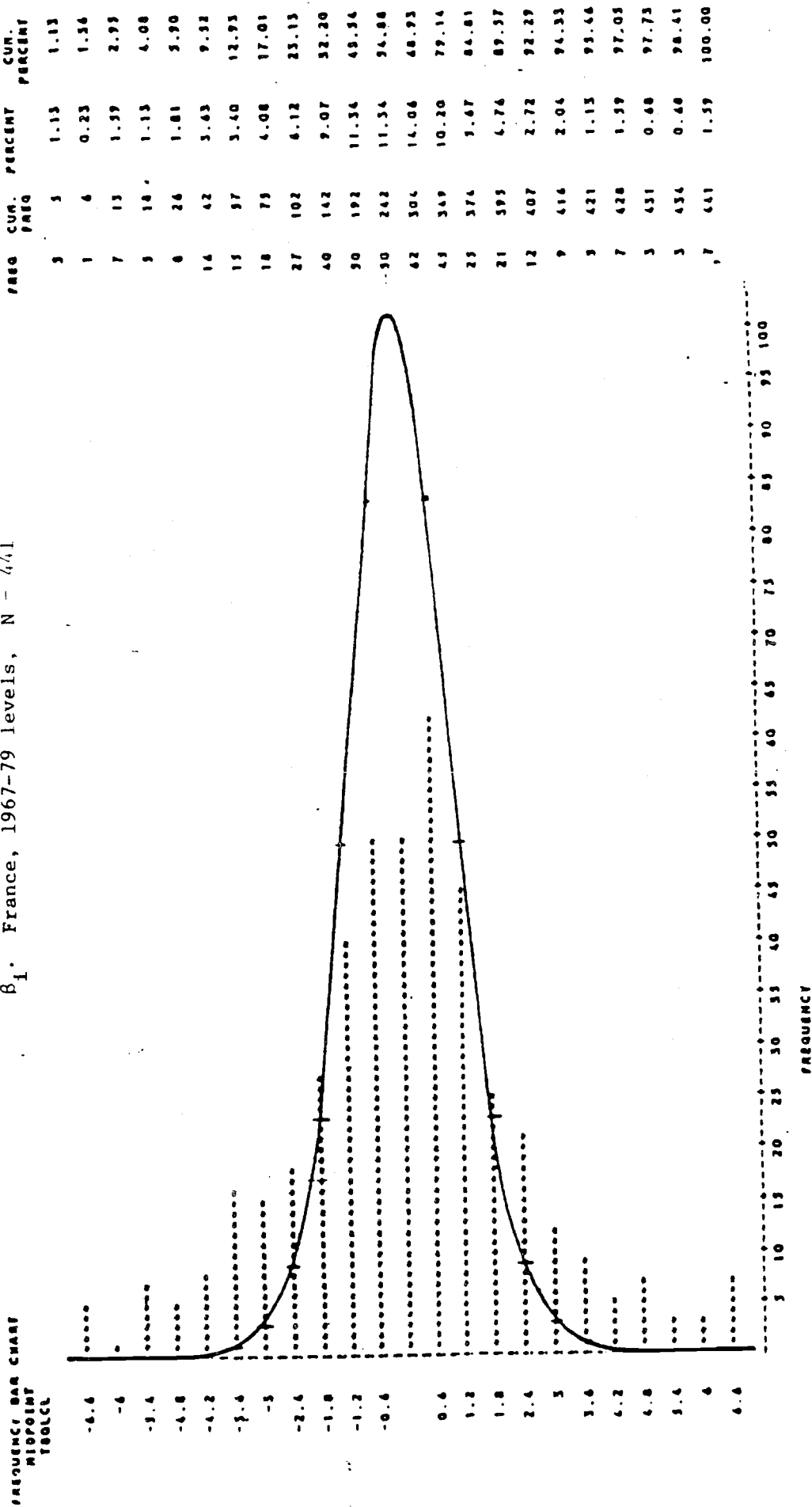


Table 1: Alternative Estimates of Production Function Coefficients

Regressions		Dependent Variable: Log Q/L					
		Levels			First Differences		
		Total	Between	Within	Total	Between	Within
FRANCE 1967-79	β	0,303 (0,009)	0,313 (0,031)	0,196 (0,011)	0,260 (0,014)	0,163 (0,039)	0,266 (0,014)
N = 441 T = 13	$\hat{\sigma}$	0,362	0,338	0,133	0,122	0,024	0,119
	R^2	0,174	0,192	0,052	0,064	0,038	0,066
JAPAN 1967-79	β	0,452 (0,007)	0,469 (0,023)	0,278 (0,009)	0,183 (0,010)	0,359 (0,029)	0,170 (0,010)
N = 845 T = 13	$\hat{\sigma}$	0,375	0,343	0,148	0,118	0,026	0,115
	R^2	0,292	0,326	0,082	0,031	0,151	0,026
UNITED STATES 1967-79	β	0,221 (0,007)	0,222 (0,024)	0,213 (0,008)	0,289 (0,009)	0,178 (0,030)	0,294 (0,009)
N = 462 T = 13	$\hat{\sigma}$	0,286	0,266	0,106	0,099	0,019	0,097
	R^2	0,154	0,163	0,096	0,149	0,073	0,153

Log Q/L = Logarithm of output per employee.

β = coefficient of log C/L. C/L = capital per employee.

Estimated standard error in parentheses.

$\hat{\sigma}$ = standard deviation of estimated residuals.

All equations contain also individual industry level year constants.

Table 2: Variability: Different Indicators

Country and Indicator	Variables				β Coefficients		
	Log Output per Employee		Log Capital per Employee		Capital Share		
	Levels	Growth Rates	Levels	Growth Rates		Levels	First Differences
<u>FRANCE</u> (N = 441)							
Average	16.9 ¹	.045	10.5 ¹	.046	.256	.194	.232
s_b Observed dispersion	.375	.025	.526	.030	.096	.542	.414
"True" dispersion σ_β	.373	0*	.524	0*	.094	.440	.213
F test	90.7	.4	132.1	.7	21.2	2.9	1.4
<u>JAPAN</u> (N = 845)							
Average	29.5 ¹	.087	14.7 ¹	.107	.316	.273	.161
s_b Observed dispersion	.418	.028	.508	.031	.103	.625	.434
"True" dispersion σ_β	.415	0*	.506	0*	.098	.539	.258
F test	87.6	.6	122.3	.8	11.1	3.9	1.5
<u>UNITED STATES</u> (N = 462)							
Average	29.8 ¹	.021	16.2 ¹	.051	-	.219	.283
s_b Observed dispersion	.291	.020	.528	.030	-	.403	.321
"True" dispersion σ_β	.289	0*	.526	0*	-	.323	
F test	81.5	.4	126.5	.5	-	2.8	1.9

Notes to Table 2

- The averages are the simple arithmetic means, except for the levels of the output per employee and of the capital per employee where they are the geometric means in thousand dollars per employee in 1970 prices.

Thus the averages of the individual estimates of the capital

coefficients $\hat{\beta}_i$ are simply $\bar{\beta} = \frac{1}{N} \sum_i \hat{\beta}_i$, and that of the variables x_{it} :

$$x_{..} = \frac{1}{NT} \sum_i \sum_t x_{it} = \frac{1}{N} \sum_i x_i.$$

- The observed dispersions s_b are the standard deviations between firms, i.e.:

$$s_B^2 = \frac{1}{(N-1)} \sum_i (\hat{\beta}_i - \bar{\beta})^2$$

or

$$s_B^2 = \frac{1}{N-1} \sum_i (x_{i.} - x_{..})^2.$$

- The true dispersions are computed by adjusting the observed dispersions for sampling variability. For the capital coefficients, they are the Swamy estimates:

$$\sigma_\beta^2 = s_B^2 - \frac{1}{N} \sum_i v_i^2, \text{ where the sampling variance } v_i^2 = s_{e_i}^2 / \sum_t (x_{it} - x_{i.})^2,$$

$s_{e_i}^2$ being the estimated variance of the errors in the individual regressions. For the variables, we have similarly:

$$\sigma_\beta^2 = s_b^2 - \frac{1}{T} s_w^2, \text{ where } s_w^2 = \frac{1}{N(T-1)} \sum_i \sum_t (x_{it} - x_{i.})^2$$

See Mairesse, 1988, for more detail.

With this notation, it can be shown that the F tests can be computed

simply as: $F = s_b^2 / (s_b^2 - \sigma_\beta^2)$.

Table 3: Maximum Likelihood Estimates on Levels

Specifications	L(ML)	β	σ_{ϵ}^2	σ_{α}^2	$\sigma_{\alpha\beta}$	σ_{β}^2
<u>FRANCE</u>						
(A) $\alpha = Ct, \beta = Ct$	2949.2	0.303 (0.009) (0.032)	0.1315 (0.0025) (0.0097)	--	--	--
(B) $\alpha = Rd, \beta = Ct$	7511.2	0.210 (0.011) (0.028)	0.0192 (0.0004) (0.0011)	0.114 (0.008) (0.010)	--	--
(C) $\alpha = Rd, \beta = Rd$	7890.2	0.212 (0.022) (0.022)	0.0148 (0.0003) (0.0007)	0.123 (0.009) (0.013)	-0.018 (0.008) (0.009)	0.145 (0.014) (0.019)
(D) $\alpha = Fs, \beta = Ct$	8702.3	0.196	0.0192	[0.118]	--	--
(E) $\alpha = Fs, \beta = Fs$	9732.1	[0.194]	[0.0146]	[0.194]	[-0.054]	[0.294]
<u>JAPAN</u>						
(A) $\alpha = Ct, \beta = Ct$	5290.9	0.452 (0.007) (0.023)	0.1404 (0.0019) (0.0065)	--	--	--
(B) $\alpha = Rd, \beta = Ct$	13255.8	0.303 (0.009) (0.021)	0.0239 (0.0003) (0.0008)	0.121 (0.006) (0.007)	--	--
(C) $\alpha = Rd, \beta = Rd$	14112.7	0.291 (0.019) (0.019)	0.0176 (0.0003) (0.0006)	0.135 (0.007) (0.008)	-0.010 (0.007) (0.008)	0.213 (0.014) (0.016)
(D) $\alpha = Fs, \beta = Ct$	15476.4	0.278	0.0238	[0.118]	--	--
(E) $\alpha = Fs, \beta = Fs$	17700.3	[0.273]	[0.0173]	[0.207]	[0.020]	[0.391]
<u>UNITED STATES</u>						
(A) $\alpha = Ct, \beta = Ct$	4507.1	0.221 (0.007) (0.034)	0.0820 (0.0015) (0.0084)	--	--	--
(B) $\alpha = Rd, \beta = Ct$	9232.2	0.214 (0.008) (0.019)	0.0122 (0.0002) (0.0007)	0.069 (0.005) (0.008)	--	--
(C) $\alpha = Rd, \beta = Rd$	9530.2	0.225 (0.016) (0.016)	0.0100 (0.0002) (0.0006)	0.068 (0.005) (0.008)	-0.001 (0.004) (0.005)	0.069 (0.007) (0.008)
(D) $\alpha = Fs, \beta = Ct$	10468.0	0.213	0.0122	[0.071]	--	--
(E) $\alpha = Fs, \beta = Fs$	11374.1	[0.219]	0.0099	[0.122]	[-0.023]	[0.163]

Notes to Table 3

Ct -- constant; Rd -- random; Fs -- firm specific (fixed effects).

() -- first line: usual standard errors obtained from the maximum likelihood method.

() -- second line: "robust" standard errors obtained from the pseudo-maximum likelihood method (White, 1980 and 1982; Gourieroux, Trognon, 1984).

Table 4: Maximum Likelihood Estimates in First Differences

Specifications	L(ML)	β	σ_{ϵ}^2	σ_{β}^2
(a) $\beta = Ct$	8506.7	0.260 (0.014) (0.022)	0.0148 (0.0003) (0.0007)	--
(b) $\beta = Rd$	8577.6	0.246 (0.019) (0.019)	0.0137 (0.0003) (0.0007)	0.064 (0.010) (0.010)
(c) $\beta = Fs$	8904.5	[0.232]	[0.0139]	[0.172]
<u>JAPAN</u>				
(a) $\beta = Ct$	16588.9	0.183 (0.010) (0.015)	0.0140 (0.0002) (0.0005)	--
(b) $\beta = Rd$	16684.2	0.168 (0.014) (0.014)	0.0131 (0.0002) (0.0004)	0.061 (0.007) (0.007)
(c) $\beta = Fs$	17308.4	[0.161]	[0.0132]	[0.188]
<u>UNITED STATES</u>				
(a) $\beta = Ct$	10071.5	0.289 (0.009) (0.018)	0.0097 (0.0002) (0.0004)	--
(b) $\beta = Rd$	10186.1	0.300 (0.014) (0.014)	0.0088 (0.0002) (0.0004)	0.040 (0.005) (0.005)
(c) $\beta = Fs$	10554.0	[0.283]	[0.0089]	[0.103]

Table 5: Likelihood Ratio Tests of Homogeneity versus Random or Fixed Coefficients Heterogeneity

Tests	FRANCE		JAPAN		UNITED STATES		$\chi_{0.01}^2(p)/p =$
	p	2LR/p	p	2LR/p	p	2LR/p	$F_{0.01}(p, \infty)$
(A) if (B): $\sigma_{\alpha}^2 = 0$	1	9124.0	1	15929.8	1	9450.2	6.5
(B) if (E): $\sigma_{\alpha}^2 = \sigma_{\beta}^2 = 0$	2	4941.0	2	8821.8	2	5023.2	4.6
(B) if (E): $\sigma_{\beta}^2 = 0, \sigma_{\alpha}^2 \neq 0$	1	758.0	1	1713.8	1	596.2	6.6
(A) if (D): $\alpha_i = \alpha$	440	26.2	844	24.1	461	25.9	7.1
(A) if (E): $\alpha_i = \alpha, \beta_i = \beta$	880	15.4	1688	14.7	922	14.8	7.1
(D) if (E): $\beta_i = \beta, \alpha_i \neq \alpha$	440	4.7	844	5.3	461	3.9	7.1
(a) if (b): $\sigma_{\beta}^2 = 0$	1	141.8	1	190.6	1	229.2	6.6
(a) if (c): $\beta_i = \beta$	440	1.8	844	1.7	461	2.1	7.1

LR - likelihood ratio. p - degrees of freedom.

Table 6: Estimates from Systems of Regressions of Squares and Cross Products of Residuals, Fully Constrained

	σ_{ϵ}^2	σ_{β}^2	$\sigma_{\alpha\beta}$	σ_{α}^2
FR Levels	0.003	0.042 (0.003)	0.001 (0.003)	0.078 (0.007)
JA Levels	0.002	0.062 (0.002)	-0.009 (0.003)	0.084 (0.006)
U.S. Levels	0.002	0.052 (0.002)	-0.001 (0.002)	0.030 (0.004)
FR First Dif.	0.007 (0.000)	0.051 (0.003)	--	--
JA First Dif.	0.008 (0.000)	0.041 (0.003)	--	--
U.S. First Dif.	0.004 (0.000)	0.032 (0.002)	--	--

σ_{ϵ}^2 -- for levels computed as the difference between the intercepts of the square and cross-product coefficients. For first differences, it is an estimate of $\sigma^2_{(\epsilon_t - \epsilon_{t-1})}$.

Table 7: Estimates from System of Regressions on Squares and Cross Products of Residuals, Unconstrained Across Groups of Equations

Coefficients from Equations of	FR Levels		JA Levels		U.S. Levels		FR FD	JA FD	U.S. FD
	σ_{β}^2	$\sigma_{\alpha\beta}$	σ_{β}^2	$\sigma_{\alpha\beta}$	σ_{β}^2	$\sigma_{\alpha\beta}$	σ_{β}^2	σ_{β}^2	σ_{β}^2
SQ	0.037	0.003	0.055	-0.009	0.056	-0.012	0.117	0.106	0.082
CP1	0.036	0.003	0.055	-0.009	0.056	-0.002	0.083	0.027	0.047
CP2	0.036	0.002	0.055	-0.009	0.055	-0.002	0.022	0.006	0.048
CP3	0.035	0.002	0.056	-0.009	0.055	-0.002	0.031	0.024	-0.011
CP4	0.035	0.002	0.056	-0.009	0.054	-0.002	0.042	-0.005	-0.006
CP5	0.034	0.002	0.056	-0.009	0.052	-0.002	0.031	0.035	0.003
CP6	0.035	0.001	0.055	-0.010	0.050	-0.002	-0.041	0.046	0.006
CP7	0.034	0.000	0.055	-0.010	0.048	-0.002	-0.015	-0.004	0.020
CP8	0.033	-0.000	0.055	-0.010	0.045	-0.002	-0.003	0.050	-0.008
CP9	0.033	-0.001	0.055	-0.010	0.042	-0.002	0.070	0.008	0.048
CP10	0.034	-0.001	0.054	-0.010	0.039	-0.002	0.101	-0.010	-0.014
CP11	0.032	-0.001	0.052	-0.010	0.035	-0.002	0.095	0.099	0.017
CP12	0.030	-0.003	0.050	-0.010	0.033	-0.002	--	--	--

Dependent variables:

SQ -- squared residuals; CPh -- cross-products of residuals h periods apart.

Table 8: Estimates from Regressions on Squares and Four Year Apart Cross-products of Residuals, Unconstrained Across Equations (OLS Estimates)

	Levels			Chamberlain's II			First Differences		
	FR	JA	U.S.	FR	JA	U.S.	FR	JA	U.S.
1979	.035	.034	.127				.107	.124	.104
1978	.042	.033	.149				.076	.363	.070
1977	.035	.036	.184				.149	.094	.185
1976	.053	.024	.154				.182	.103	.202
1975	.033	.032	.191				.183	.099	.170
1974	.055	.063	.157				.490	.080	.155
1973	.019	.045	.170				.231	.136	.114
1972	.021	.046	.163				.186	.128	.089
1971	.034	.048	.150				.129	.091	.146
1970	.017	.043	.136				.154	.135	.105
1969	.006	.026	.078				.125	.085	.087
1968	.009	.042	.059				.095	.136	.078
1967	.015	.035	.027				--	--	--
1979-75	.040	.034	.165				-.178	.019	-.089
1978-74	.045	.045	.160				.163	-.050	.005
1977-73	.027	.043	.185				.098	.056	-.071
1976-72	.036	.043	.172				.056	-.012	.180
1975-71	.038	.046	.171				-.033	-.107	-.023
1974-70	.031	.059	.145				-.001	.044	-.103
1973-69	.011	.042	.122				.067	.062	.040
1972-68	.020	.055	.099				.072	-.074	.019
1971-67	.021	.053	.073				--	--	--

Table 9: The Instability of Heterogeneity Estimates
1967-79 versus 1967-73 and 1974-79

Country and	Labor productivity	Capital Share	Capital Coefficients (β)	
			Levels	First Differences
<u>France</u>				
s_b observed	.375	.096	.542	.414
σ_β "true" (tp)	.373	.094	.440	.213
σ_β "true" (bp)	.365	.089	.110	.113
Corr bp	.90	.76	.05	.05
Corr ($s, \hat{\beta}$)			.04	-.01
<u>Japan</u>				
s_b observed	.418	.103	.625	.434
σ_β "true" (tp)	.415	.098	.539	.258
σ_β "true" (bp)	.405	.094	.186	.180
Corr bp	.88	.68	.06	.09
Corr ($s, \hat{\beta}$)			-.01	.02
<u>U.S.</u>				
s_b observed	.291		.403	.321
σ_β "true" (tp)	.289		.323	.218
σ_β "true" (bp)	.282		.152	.154
Corr bp			.09	.12

s_b observed -- computed from the dispersion of the separately estimated coefficients (or the firm variable means). From Table 2.

σ_β "true" (tp) -- estimated "true" dispersion, based on the total period 1967-79 (tp). From Table 2.

σ_β "true" (bp) -- estimated from the covariance between the separately estimated capital coefficients in the two periods (or the means of the variables).

Corr bp -- observed correlation between the separately estimated capital coefficients in the two periods.

Corr (sc, β) -- correlation between the average capital share and the estimated (total period) capital coefficients. Unavailable for the U.S.

Table 10: Summary of Heterogeneity of β_i Estimates for France

	Levels		First Differences	
	σ_β	σ_β^2	σ_β	σ_β^2
0. Observed	.542	.293	.414	.171
1. Swamy (tp)	.440	.194	.213	.045
2. Maximum Likelihood	.382	.014	.253	.064
3. SQ and CP (total)	.205	.042	.226	.051
4. π residuals SQ and CP	.187	.035	n.r.	n.r.
5. π residuals "distant" CP	.119	.014	n.r.	n.r.
6. Between periods	.110	.012	.113	.013

1. From Table 2.

2. From Tables 3 and 4.

3. From Table 6.

4. and 5. Computed from residuals of regressions in which each y_{it} was regressed on all (13) x 's.

5. Based only on regressions of cross-products 10 years or more apart.

6. From Table 9.

n.r. -- not relevant.

Table 11: Alternative Estimates of the Capital-Output Elasticity β

Levels	France	Japan	United States
1. Between	.313 (.031)	.469 (.023)	.222 (.024)
2. Within	.196 (.011)	.278 (.009)	.213 (.008)
3. Average of Individual β 's	.194 (.026)	.273 (.022)	.219 (.019)
4. ML Random Effects $\bar{\beta}$.212 (.022)	.291 (.019)	.225 (.016)
<u>First Difference</u>			
5. Average of Individual β 's	.232 (.020)	.161 (.015)	.283 (.015)
6. ML Random Effects $\bar{\beta}$.246 (.019)	.168 (.014)	.300 (.014)

Lines 1 and 2 from Table 1, lines 3-6 from Table 3. Standard errors for lines 3 and 5 computed from the observed variance across firms divided by $N-1$.

Table 12: Alternative Swamy's Estimates for France

Various Types of Estimates	Estimates in Levels			Estimates in First Differences			Estimates in Quasi Differences		
	β	σ_α^2	$\sigma_{\alpha\beta}^2$	β	σ_β^2	σ_β^2	β	σ_β^2	σ_β^2
1. Observed	0.194	0.194	-0.054	0.293	0.171	0.232	0.171	0.227	0.187
2. "True" if $\Sigma_i = \sigma_{ei}^2 I$	0.212 (0.024)	0.167	-0.050	0.181	0.011	0.254 (0.015)	0.011	0.227 (0.018)	0.072
3. "True" if $\Sigma_i = \sigma_{ei}^2 I$	0.196 (0.024)	0.171	-0.048	0.193	0.045	0.233 (0.016)	0.045	0.220 (0.018)	0.089
4. "True" if $\Sigma_i = \text{diag}(\hat{\Sigma})$	0.211 (0.024)	0.171	-0.050	0.195	0.023	0.251 (0.015)	0.023	0.229 (0.018)	0.078
5. "True" if $\Sigma_i = \sigma_{\epsilon}^2 \text{TOOR}$	0.219 (0.023)	0.152	-0.050	0.132	0.011	0.248 (0.014)	0.011	—	—

Quasi-differences: $Y_t - .5Y_{t-1}, x_t - .5x_{t-1}$.

Σ_i — variance-covariance matrix of the residuals for firm i .

$\sigma_{\epsilon}^2 \text{TOOR}$ — σ_{ϵ}^2 multiplied by a matrix which corresponds to $\rho_r = \rho^T$, and $\rho = .5$ assumption.