LONGITUDINAL ECONOMIC DATAAT THE CENSUS BUREAU: A NEW DATABASE YIELDS FRESH INSIGHTSON SOME OLD ISSUES

by

Robert H. McGuckin*

CES 90-1 January 1990

Abstract

This paper has two goals. First, it illustrates the importance of panel data with examples taken from research in progress using the U.S. Census Bureau's Longitudinal Research Database (LRD). Although the LRD is not the result of a "true" longitudinal survey, it provides both balanced and unbalanced panel data sets for establishments, firms, and lines of business. The second goal is to integrate the results of recent research with the LRD and to draw conclusions about the importance of longitudinal microdata for econometric research and time series analysis. The advantages of panel data arise from both the micro and time series aspects of the observations. This also leads us to consider why panel data are necessary to understand and interpret the time series behavior of aggregate statistics produced in cross-section establishment surveys and censuses. We find that typical homogeneity assumptions are likely to be inappropriate in a wide variety of applications. In particular, the industry in which an establishment is located, the ownership of the establishment, and the existence of the establishment (births and deaths) are endogenous variables that cannot simply be taken as time invariant fixed effects in econometric modeling.

*Robert H. McGuckin, Chief, Center for Economic Studies, (CES) Bureau of the Census, Washington, DC 20233. This paper does not necessarily reflect the opinions of the Census Bureau. The author acknowledges the comments and many useful discussions with Timothy Dunne and John Haltiwanger. Helpful comments from Stephen Andrews, Scott Schuh, James Monahan, Sang Nguyen, and Mary Streitwieser are appreciated. Stephen Andrews and Robert Bechtold did the data preparation and Janice Bryant and Terry Norris provided excellent typing services.

I. <u>Introduction</u>

"You can't always get what you want, but if you try sometime ... you get what you need." (Let It Bleed, 1969, Mick Jagger and Keith Richards)

This paper has two goals. First, it illustrates the importance of panel data with examples taken from research in progress using the U.S. Census Bureau's Longitudinal Research Database (LRD). A panel data set is one that contains multiple observations on economic entities over time. For example, an establishment panel data set might have observations on shipments across individual plants linked over time. In contrast, time series data usually refer to observations over time on an aggregate economic variable, such as total industry shipments or U.S. national income. The advantages of panel data arise from both the micro and time series aspects of the observations.

Although the LRD is not the result of a "true" longitudinal survey, it provides both balanced and unbalanced panel data sets for establishments, firms, and lines of business.[1] The LRD enables researchers to conduct many essential studies heretofore considered impossible. In this sense, "you get what you need."

The second goal is to integrate the results of recent research with the LRD and to draw conclusions about the

---

[1]The committee on Statistical Methodology argued that the essential feature of a longitudinal survey is that "from the beginning, there is a plan to elicit data from the future for each observational unit." (1986) The committee contrasted longitudinal surveys with surveys that support longitudinal analysis. The Longitudinal Research Database (LRD) was put, correctly, in the latter grouping. It was the only establishment panel among the 12 data sets studied.

importance of longitudinal microdata for econometric research and time series analysis. The discussion focuses on research involving the behavior of firms and establishments. This also leads us to consider why panel data are necessary to understand and interpret the time series behavior of aggregate statistics produced in cross-section establishment surveys and censuses.

Most economic modeling is based on theories concerning the behavior of individual economic agents. Estimation and inference based on aggregate data involve assumptions about the homogeneity of the individual entities making up the aggregate. For example, a typical assumption might be that the distribution of the entities with respect to a particular variable such as efficiency or industry classification remains constant over time. This study indicates that typical homogeneity assumptions are likely to be inappropriate in a wide variety of applications. The evidence illustrates a basic point: The industry in which an establishment is located, the ownership of the establishment, and the existence of the establishment (births and deaths) are endogenous variables which cannot simply be taken as time invariant fixed effects in econometric modeling.[2]

[2]Most of the work with panel data has relied on data and models based on individuals. While many of the techniques applicable to individuals can be carried over to models of the behavior of firms and establishments, some new issues are involved. For example, the importance of ownership changes to establishment behavior has no obvious analogue for the

The importance of panel data sets for economic research cannot be overestimated. Many economic issues simply cannot be addressed in the absence of panel data. As noted, these issues include a wide range of questions involving behavior before and after particular policy actions or other changes in the circumstances or environment of economic agents. Panel data sets also provide a unique vehicle for calculating microlevel measures of gross changes that are often missed in the aggregate statistics.[3]

New evidence from CES research suggests that measures of gross change are important for many issues which have generally been examined with data on net changes. For example, some new work on job turnover finds that gross job reallocations are important in both a time-series (business cycle) and cross-section (across establishments and industries) sense. Work dealing with the entry and exit of firms and plants reaches similar conclusions. Analysis of one measure of turnover in industrial markets is used to contrast the importance of gross with net flow measures commonly available for analysis.

While the mechanisms at work are not completely understood,

---

individual. Analogies to the household as a unit of analysis are likely to be most apt. Nonetheless, the analogies are not perfect.

[3]Statisticians also emphasize the use of panel data in reducing collinearity and improving the precision of estimates in dynamic economic models involving lagged explanatory variables.

there are several reasons for expecting gross change measures to have important economic impacts. First, change typically requires resources and therefore measures of gross change provide a basis for measuring and understanding such costs. Second, the evidence that ownership change affects performance suggests that gross turnover measures provide important information on competitiveness.

A third reason for examining gross changes is that they provide a basis for determining if aggregate movements are being generated by a large or small segment of economic entities. Knowing how broadly based are the forces behind aggregate movements is important for policy makers. Longitudinal panels are required to address the issues involved in each of these examples.

Because the LRD is relatively new, a brief description adds some concreteness to the discussion. It also provides a basis for evaluating the LRD as a source of panel data.

II. The LRD

The LRD is constructed by linking together individual establishment records from the Census of Manufactures (CM), which takes place every 5 years, and the Annual Survey of Manufactures (ASM), conducted each year. At present, the LRD has substantially more than 2 million manufacturing establishment-year records including information on over 800,000 different establishments in the 1963-86 period. When the 1987 census is included in the database, the number of unique establishments will likely jump to over 1 million.

Table 1 provides a tabulation of the number of establishments in the LRD in each year. Each census year, 1963, 1967, 1972, 1977, and 1982 contains well over 300,000 establishments of which about two thirds are actually surveyed. The administrative record cases, those which are not directly surveyed, represent small establishments (primarily establishments with less than 5 employees) which have little impact on aggregate industry totals. In non-census years the LRD contains roughly 70,000 establishments in the period 1973-78 and 55,000 after 1979 when there was a major redesign of the ASM.

The probability that any plant is sampled for the ASM is directly related to its size. However, the relationship is complicated. Large establishments, those with more than 250 employees, are sampled with certainty. Among the remaining smaller establishments (those with employment less than 250 and greater than 10), establishments are sampled with probabilities directly related to employment size except that there is an attempt to exclude those establishments sampled in one panel in the following panel.[4] This rotating panel design is to reduce the reporting burden on small plants. New panels are chosen

---

[4]The 1984 and new 1989 panels include with certainty the largest 500 firms in 1984 and 1989.

every 5 years with the primary aim of obtaining accurate estimates of aggregate industry variables such as shipments.[5]

[5]While additions and subtractions to the sample are made each year to account for the formation of new establishments and the closing of existing establishments, there are various lags in the process, and some uncertainties are not resolved until census years.

Cross-Section Design

The LRD data are collected from surveys and censuses that are cross-sectional in design and processing. While the processing procedures include previous year values in the edit sequences, there are few time based edits. As an example of the cross-section design, some large establishment reports are split into two or more establishments when the establishment produces a variety of distinct outputs. This procedure increases the precision of industry aggregates in the cross-section, but reduces the accuracy of establishment linkages across time by making it more difficult to trace individual plants.

The rotating panel design and the fact that most establishments are not sampled for the ASM do not in principle have any effect on the cross-section aggregate estimates. However, it does make following establishments over time more difficult and reduces the number of establishments who have continuous data for every year. The effects of this design on the availability of consistent yearly panels from the LRD are significant. For the 1972-86 period there are only a little more than 16,000 establishments that have data in every year in the LRD, less than 5 percent of the establishments in existence in any year.[6]

Establishments not sampled in the ASMs appear only in the

---

[6]This number represents approximately 30 percent of the total ASM sample of establishments.

CMs. With 5 censuses available and another (1987) to be available soon, the possibilities for research based on balanced panels with observations at 5-year intervals are good.[7] Information on the composition of the linkages available for the census years 1972, 1977, and 1982 is presented in Roberts and Monahan (1986). They show that of the roughly 600,000 unique establishment records identified in these 3 years, approximately 133,000, or 22 percent, are present in all 3 years. These data were extended to the 1963-82 period by Dunne and Roberts (1986). For the 1963-82 period approximately 66,000 linked establishments are available to form a balanced panel. Although attrition will reduce the panel number by 1987, there still should be over 50,000 establishments observed continuously from 1963-87.

Data

The LRD contains a variety of information on individual establishments. Most of the data are reported on a yearly basis, but employment and hours worked are provided quarterly. By and large, the data contained in the LRD relate to production and various classification and identification characteristics of establishments. The latter category includes information on the plant's ownership, location, age (for some plants), product and industry structure, and various status codes which identify, among other things, birth, death and ownership changes. These

---

[7]Even here, however, the traditional emphasis on aggregate tabulations has had an adverse effect on the available linkages.

identifying codes are used in developing both the longitudinal plant linkages and ownership linkages among plants.

Most of the data collected for each plant provides information on the inputs or outputs of the plant. A detailed description of the individual data items would be too lengthy to include here, but can be found in the LRD Technical Documentation available from CES that maintains and updates the LRD. However, the list of variables shown in Table 2 gives a good idea of the breadth of coverage. On the input side the LRD contains data on major factors of production; labor (production and other), capital, materials, and purchased services.

The output data include value of shipments reported for each 7-digit product in census years and at the 5-digit level of detail in ASM years. Related information, such as value added, miscellaneous receipts, inventories, value of resales, and receipts for contract work are also available for each establishment.

For the most part price data can be derived in census years in the form of unit values.[8] Outside of census years the quantity data to calculate unit values is not available in the LRD. This means that price series for purposes of, for example,

---

[8]Current Industrial Reports data are not linked to the LRD. These reports contain yearly and sometimes monthly unit value data for many detailed SIC classifications. The CES has several specific projects working with these data and hopes to eventually be able to link CIR data more generally to the LRD.

deflation in production function estimation must be based on industry level price series. Such a series is published by the U.S. Department of Commerce based on Bureau of Labor Statistics (BLS) data. This series has been used by several researchers for purposes of deflation.[9]

III. LRD Research and Time

The research program at CES emphasizes projects that exploit the longitudinal characteristics of plants and firms. Many projects are measurement orientated. They establish important sets of "stylized facts" that form the basis for more substantive hypothesis testing. Examples of work in this category are studies by Dunne, Roberts, and Samuelson (1988, 1989b) dealing with patterns of firm entry and exit and gross employment flows, respectively. Both of these studies used 5 year panels formed from census year data in the LRD. Other work in this category is reported in Davis and Haltiwanger (1989), where new measures of gross and net employment fluctuations at yearly intervals are constructed. Other studies at CES are oriented toward testing particular hypotheses. In this category of work there are various studies examining the importance of ownership changes on plant and firm performance that exploit the longitudinal

---

[9]Some recent research suggests that prices differ across establishments and areas. Thus, the establishment may be a more appropriate level for deflation for certain research projects (see Abbott (1989)).

structure of the LRD. Examples in this category are work by McGuckin and Andrews (1988), and Lichtenberg and Siegel (1988, 1989a, 1989b).

In what follows no attempt is made to be exhaustive in describing LRD research. For example, a wide variety of work on productivity measurement is not discussed in any detail. (Several studies have been published in the last 2 years and there are several other major projects underway.) The purpose is to illustrate the types of research for which panel data such as those contained in the LRD are essential. But, even more important, is the evidence that these types of analysis are crucial to understanding important economic phenomena and making informed policy judgements.

The Behavior of a Plant Over Time

It is useful to begin with a simple model to characterize the performance or behavior of an economic entity such as a plant or firm. For concreteness, assume that the i_{th} plant's performance at time t, Y_{it}, can be described by the relationship

(1) Y_{it} = " + μ_j + 8_t + $E_s\$_s X_{sit}$ + $,_{it}$

where X_{sit} are exogenous explanatory variables, " represents plant level fixed effect common to all plants, μ_j is a time invariant fixed effect such as ownership, industry, or location which is common to a group of plants, 8_t is a time varying fixed effect that is constant over individual plants, and $,_{it}$ is an error term. This simple model of plant performance can be used to characterize the issues of interest.

One important question is what can be controlled with the fixed effects specification. Of particular interest here is the question of what are the time invariant effects which can be represented by μ_j. One obvious candidate is the industry classification of the plant. Another is the ownership of the plant. Neither candidate is satisfactory.

Ownership Changes

The size and scope of the recent merger movement makes clear that plant ownership is often changed. Treating ownership characteristics as time-invariant is appropriate if the plant's behavior remains relatively unchanged before and after the ownership change. But, studies with the LRD indicate that ownership changes have dramatic effects on operating performance, whether measured at the plant or firm level.

McGuckin and Andrews (1988), find that line of business market shares increase relative to those of lines of business not experiencing a merger, particularly for complete firm takeovers. Lichtenberg and Siegel (1988, 1989b) find improved plant productivity following ownership change. Furthermore, they are able to associate most of this gain with fewer administrative employees and lower wages for them following ownership changes (Lichtenberg and Siegel (1989a)). These kinds of "event" studies are impossible without a panel of observations on individual establishments or firms.[10]

Primary Industry Affiliation

Various work has also shown that the industry category of establishments changes frequently. Approximately one third of the panel of more than 16,000 establishments continuously observed from 1972-86 experienced a switch in their primary 4-digit industry. Thus, treating industry as a fixed effect may be a mispecification of the model.

The balanced LRD panel is generally over-represented by large establishments. Thus, it is not simply small plants with little total output that are involved in industry classification switches. Abbott and Andrews (1988), report that primary 4-digit industry switches among plants in contiguous censuses account for over 3 percent of total output in both the 1972-77 and 1977-82 periods. For some 2-digit classifications, the average 4-digit industry had 10 percent of its output involved in switches. In short, the output effects of these switches on industry totals are significant in many industries.

Product Class Affiliations

These observed switches, in addition, are not simply the result of measurement errors associated with multiple output plants being reclassified from one industry to another because

[10] Economic studies of this type are not widespread. One of the few areas where such work is common is in the finance literature. See McGuckin, Warren-Boulton, and Waldstein (1988) for an example of an "event" study using stock market data.

the "primary" output of the plant changes. Much of the "switching" activity involves adding or dropping whole product areas. Based on a comparison of matched establishments observed in both the 1981 ASM and the 1982 census, we found, based on some data developed as part of a study by the Department of Commerce's Bureau of Economic Analysis (BEA), that the gross outputs involved in switches averaged over 10 percent of total output in 1981 for both switches into and out of a product class.[11]

Tables 3 and 4 provide data on the percentage of product class output produced by plants that had production in the product class in 1981, did not produce in the product class in 1982, but did produce in one or more different product classes in 1982. This figure is termed the percentage of output that "switched out" of the product class. "Switched in" output is analogously defined also using the 1981 product class output as a base.

The Tables show that the distribution of the gross changes are dispersed and quite large. The average product class had over 10 percent of its output switched. In roughly 75 percent of the product classes, gross output attributable to switches in or out is more than 5 percent. In 5 percent of the product classes, over 70 percent of the output represented switches.

---

[11]For various reasons associated with the ASM sampling design, comparisons based on census and immediately preceding ASM years will overstate the annual switching rate. Even so, these numbers are large.

The net output effects of switches are substantially smaller, averaging less than 3 percent of total 1981 product class output. But as shown in Table 5, the distribution of the net changes shows relatively high values in certain product classes. The Table indicates that in over 10 percent of the roughly 1,350 usable product classes, the net effect of switches by matched establishments is greater than 5 percent.[12]

The data in Tables 3-5 reflect all 5-digit product groups available for analysis. One might object that this procedure overstates the problems by including a variety of miscellaneous product classes. This is indeed true for the net changes shown in Table 5. When we excluded all product classes ending in zero or 9, the miscellaneous categories, over 90 percent of the product classes showed the percentage of 1981 total output represented by net switches in the range between -3.5 and +3.5 percent. Nonetheless, for this set of product classes the percentage of total output subject to switches, although smaller than found in Tables 3 and 4, still averaged around 9 percent with a standard deviation of about 15. The phenomenon of large proportions gross output being associated with switches is not

[12]These figures are calculated from data for the roughly 50,000 establishments sampled in the ASM panel. The output totals in 1981 are the product class totals published by BEA. The 1,337 product categories used in this study included all those with complete data and comparable definitions for each year. About 200 product classes were eliminated in the edit process.

simply the result of poorly defined product classes.

These findings suggest that industry effects cannot be simply thought of as time-invariant effects. Moreover, switches are not simply the result of random (or nonrandom) measurement errors arising from problems with the SIC classification system. Research has not yet established if the probability of switches is greater at the time of ownership change than at other points in an establishments' history. Some evidence (McGuckin and Andrews (1988)) points this way. But managements do make economic decisions to reallocate a plant's productive capacity to new activities. This is true at the time of ownership changes. It is also true in day to day decisions as multiple product plants shift production in response to, among other things, changes in product demand. This suggests that for some problems at least, switches need to be treated as endogenous or an explained phenomenon.[13]

Implications for Aggregate Time Series Data

It is important to recognize that the work cited above has implications for aggregate analysis. The first implication is directly related to the discussion of the applicability of models of the type represented by equation (1). Aggregate analysis makes use of assumptions concerning the nature and homogeneity of

---

[13]In the merger studies noted above many industry and product class switches are associated with establishment ownership changes.

individual economic agent's behavior. The evidence from LRD based research suggests that typical assumptions about the "representativeness" of aggregate observations may be inappropriate.

Aside from the modeling aspect, a second implication concerns the character of the observed time series. Because of various processing considerations, most changes in industry output (or employment) associated with switches occurs in census years. One of the primary reasons for this is that as part of each census, a complete canvass of establishments is undertaken and an extensive company organization survey is conducted. Firms are asked to give a description of all products produced in their plants. The Census Bureau uses the new information to reclassify plants, so that plants are sent correct survey forms for the census. It is thus in census years that many of the switches are identified.[14]

The large portions of industry output that are subject to switches and the realities of processing imply that published aggregate output, employment, or other establishment based variables will contain discrete jumps between ASM and census years. Observation of these jumps led to the BEA project that collected the data underlying Tables 3 to 5. Recognition of the

---

[14]Moreover, the ASM sampling design has resistance rules that limit establishment reclassifications in non-census years. Also, entrants are very difficult to track down and are often not observed directly until the organization survey is completed.

source of these jumps should provide information to improve the quality of available time series. It also should aid in the development of reconstructed time series that can be compared to unadjusted time series data obtained from the traditional cross-section aggregations of surveys and censuses. Such studies should yield important information for the interpretation of time series models.

Gross Flows

The LRD enables one to develop information on gross as well as net flows of economic variables such as job creation and entry. The opportunities for examination of measures of gross change has motivated a number of studies at CES. These studies collectively suggest that reliance on aggregate cross-sectional measures of net change may obscure important economic phenomena.

Job Reallocations

Recent work by Davis and Haltiwanger (1989) suggests that gross measures of job creation are important in the study of business cycles and other macroeconomic issues. Davis and Haltiwanger find that manufacturing employment contracted at a rate substantially less than the rate of gross job reallocations (the sum of job creation and destruction rates) in the 1972-86 period. The size of the gross reallocation rates relative to the observed net charges, (roughly 10 percent points greater) implies the existence of large worker flows across establishments that are masked by examination of net changes. Further, Davis and

Haltiwanger find that gross job reallocation exhibits significant countercyclic time variation in contrast to the procyclical behavior of net job reallocations. The important point that Davis and Haltiwanger make is that gross measures of job creation and job destruction are important in the study of business cycles and other macroeconomic issues. As illustrated next, gross flow measures are also important in examining microeconomic issues.

Entry and Exit

The importance of the structure of a market in determining performance has long been emphasized. Until fairly recently, studies often relied on measures of market structure such as concentration ratios as an indicator of the likelihood of monopoly power. A concentration ratio measures the share of output produced by, for example, the largest four firms in a market.[15] In its simplest form, the theory suggests that the concentration ratio provides a measure of the ease of coordinating pricing policies by the largest firms in an industry.

There are many problems in using a concentration ratio alone as a measure of monopoly power. Among the most important is the long recognized importance of entry (or potential entry) as the ultimate constraint on firms that price above competitive levels.

---

[15]The choice of the four-firm ratio rather than the three or two-firm ration has its origins in the confidentiality protection rules employed by the Census Bureau.

Until recently, little information has been available to construct measures of entry beyond simple net changes in numbers of firms.

One possibility for creating a dynamic measure of market structure based on gross entry and exit is to measure the number of large firms in a market who survive from one point in time to another. The theoretical justification for this measure is that it captures information about the turnover of competitors in a market. The measure is not new and the empirical tests reported here are only suggestive. But, they illustrate the possibilities and importance of longitudinal considerations in examining market structure.

The survival measure is developed for the roughly 450 4-digit industries in manufacturing for the years 1972-77, 1977-82, and 1972-82. The actual calculations included the 20 largest firms in terms of value of shipments in each census year in each industry. Thus, the number of survivors is simply the converse of the gross turnover of firms over the period. That is, we measure gross turnover simply as the total number of firms (20) less those firms that remain in the top 20. By construction, net turnover is zero.

Using the top 20 producers reduces the possibilities of misclassifications of small firms. For most industries the top 20 producers account for the bulk of industry output. They account for over 60 percent of industry output in 280 of the 450 available industries. Only in 55 industries did the top producers account for less than 40 percent of output. In fact, the average industry had roughly 75 percent of its output accounted for by the largest 20 firms in the 3 census years under consideration, 1972, 1977, and 1982.

The results of the calculations showed significant turnover among the largest firms. Table 6 shows that in a time span of as little as 5 years, the average industry replaced 8-9 top 20 firms. This figure implies a gross turnover rate of approximately 40 percent (8/20) for both the 1972-77 and 1977-82 periods. Measured across the 10 year interval 1972-82, gross turnover averages almost 60 percent with 11-12 of the top 20 firms replaced in the average industry.

These turnover rates do not suggest the widespread exercise of monopoly power. But, even with this large turnover, the market shares of the largest firms may be quite stable. This is the implication from the survival rate breakdowns by concentration class in Table 6.

Turnover, as expected, is greater when measured among the 20 largest firms than when measured among the 8 or 4 largest. Thus, reading across the rows in Table 6 one always finds the percentage survival rate increases with the number of firms in the initial size distribution, top 20, top 8, or top 4. However, there is little difference in turnover rates across concentration classes. Although the percentage of survivors was always smallest for the less than .4 concentration class, there is little difference between the two largest classes and the difference between these classes and the smallest also is not large. Moreover, regardless of the initial levels of concentration in the industry, the average industry lost approximately one firm from among the top four firms at the beginning of each period.

Although direct comparisons are not possible because of differences in procedure, Dunne, Roberts, and Samuelson (1988) also develop gross entry and exit rates for 4-digit industries.[16] A major difference in procedure relates to the treatment of ownership changes. Dunne, et al do not treat firms with ownership changes as new entrants unless the change alters the basic establishment structure of the firm in the market. They only consider as entrants firms bringing new capacity to the market. If new management takes over existing plants, this is treated as a name change.

In contrast, in this paper all ownership changes are treated as entrants and exits in calculating survival rates. If a new competitor is defined in terms of the capacity it brings to the market, then excluding "name changes" resulting from a merger or other ownership change makes sense. However, if, as suggested by

---

[16]Their study includes all producers of a product, not just primary producers. They also exclude the smallest firms, those in aggregate accounting for less than one percent of total industry output.

the work on ownership changes cited earlier, new ownership brings new management and increased performance, then the "name" changes should count as entrants.

How much they should count is another question. One that cannot be decided on a *priori* grounds. Only with further empirical work relating performance and behavioral measures to survival rates and other measures of dynamic concentration will it be possible to sort out the proper measures.

The one sure conclusion that we draw from these studies is that panel data are necessary to make progress on these issues. While we have focused on the cross-section or across industry variation in turnover in this example, as with the work of Davis and Haltiwanger (1989) cited earlier, time series variations, reflecting shifts in demand, technological opportunities, or shifts in input prices, are likely to be important components of net and gross turnover. This is clearly the implication of the existence of merger cycles that have been identified with, among other things, industry shocks (see Blair (1989)).

IV. Concluding Remarks

As suggested by the quote at the beginning of the paper, the available panels in the LRD permit a wide range of longitudinal studies. Here we emphasize two generic classes of studies that can be accomplished with panel data. The first is the so-called event study. In the examples cited, we show the importance of incorporating time varying effects in explaining establishment and firm behavior (both existing and new). Various studies at CES have shown distinct differences in firm performance following ownership changes. This work suggests that ownership changes need to be incorporated in models explaining establishment and firm behavior. Moreover, since the volume of mergers and other forms of ownership change varies greatly over time, these kinds of changes can have significant effects on aggregate time series data.

In this regard we also report on the large volume of establishment industry switches. These switches can generate jumps in aggregate industry output time series since, for a variety of reasons, the effects of such switches are largely accounted for in census years. In addition, since there is some evidence that these switches arise from ownership changes and other corporate events, their effect on aggregative output measures is not simply a processing or sample design consideration. Rather, it is a phenomenon to be modeled. At the very least, given the increased, number of mergers and acquisitions observed in the 1980s, an assessment of the effects of switches on aggregate statistics needs to be undertaken. We noted that current work at CES finds that gross job turnover measures have important implications for analysis of labor markets and business cycles. In addition, the importance of measuring gross flows was illustrated with a simple "dynamic" measure of market structure which exploited the LRD to obtain a measure of gross entry and exit.[17] In this respect, as well as in the event studies, we "get what we need."

[17]While we make an attempt to discuss these here, work in Canada on a database similar to the LRD is also suggesting that gross flow measures are extremely important for analysis of competition and export productivity, and labor job reallocations. Various work by Baldwin and Gorecki (1989b, 1989c, 1989d, 1989e) suggests the Canadian Experience is very similar to the U.S. during the 1970s.

References

Andrews, Stephen H. and Thomas A. Abbott, III (1988), "An Examination of the Standard Industrial Classification of Manufacturing Activity Using the Longitudinal Research Data Base," Bureau of the Census Fourth Annual Research Conference Proceedings.

Blair, Margaret (1989), "Free Cash Flow and the Rise in Contests for Corporate Control," Brookings Discussion Paper in Economics.

Baldwin, John R. and Paul K. Gorecki (1987), "The Dynamics of Firm Turnover," Economic Council of Canada Working Paper.

Baldwin, John R. and Paul K. Gorecki (1989a), "Measures of Market Dynamics: Concentration and Mobility Statistics for the Canadian Manufacturing Sector," Forthcoming Annales D' Economic et Statistique.

Baldwin, John R. and Paul K. Gorecki (1989b), "Job Turnover in Canada's Manufacturing Sectors," Statistics Canada Analytical Studies Branch Research Paper Series, No. 22.

Baldwin, John R. and Paul K. Gorecki (1989c), "Firm Entry and Exit in the Canadian Manufacturing Sector," Statistics Canada Analytical Studies Branch Research Paper Series, No. 23.

Baldwin, John R. and Paul K. Gorecki (1989d), "Mergers Placed in the Context of Firm Turnover," Department of Economics, Queen's University, and Statistics Canada; draft.

Baldwin, John R. and Paul K. Gorecki (1989e), "Dimensions of Labor Market Change in Canada: Intersectoral Shifts, Job and Worker Turnover," Business and Labor Market Analysis Group Statistics Canada; draft.

Caves, R.E. and M.E. Porter (1978), "Market Structure, Oligopoly, and Stability of Market Shares," Journal of Industrial Economics, Vol. XXVI, pp. 289-313.

Caves, Richard E. and Michael E. Porter (1980), "The Dynamics of Changing Seller Concentration," Journal of Industrial Economics, Vol. XXIX, pp. 1-15.

Davis, Steve J. and John Haltiwanger (1989), "Gross Job Creation, Gross Job Destruction and Employment Reallocation," forthcoming in Center for Economic Studies Discussion Paper.

Dunne, Timothy and Mark J. Roberts (1986), "Measuring Firm Entry and Exit With Census of Manufacturers Data," Department of Economics, The Pennsylvania State University.
Dunne, Timothy and Mark J. Roberts (1987), "The Duration of Employment Opportunities In U.S. Manufacturing," Department of Economics, The Pennsylvania State University.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson (1988a), "The Growth and Failure of U.S. Manufacturing Plants," Department of Economics, The Pennsylvania State University Working Paper.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson (1988b), "Patterns of Firm Entry and Exit in U.S. Manufacturing Industries," The Rand Journal of Economics, Vol. 19, No. 4 (Winter).

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson (1989), "Firm Entry and Post-Entry Performance in the U.S. Chemical Industries," Center for Economic Studies Discussion Paper, CES 89-6.

Gort, Michael (1973), "Analysis of Stability and Change in Market Shares," Journal of Political Economy, Vol. 71, pp. 51-63.

Grossack, Irvin M. (1965), "Towards an Integration of Static and Dynamic Measures of Industry Concentration," Review Economics and Statistics, Vol. 47, pp. 301-308.

Lichtenberg, Frank and Donald Siegel (1987), "Productivity and Changes in Ownership of Manufacturing Plants," Brookings Papers on Economic Activity, pp. 643-673.

Lichtenberg, Frank R. and Donald Siegel (1989a), "The Effect of Takeovers on the Employment and Wages of Central-Office and Other Personnel," Center for Economic Studies Discussion Paper, CES 89-3.

Lichtenberg, Frank R. and Donald Siegel (1989b), "The Effects of Leveraged Buyouts on Productivity and Related Aspects of Firm Behavior," Center for Economic Studies Discussion Paper, CES 89-5.

McGuckin, Robert H. (1972), "Entry, Concentration Change, and Stability of Market Shares," Southern Economic Journal, Vol. XXXVIII, No. 3.

McGuckin Robert H. and Sang V. Nguyen (1988), "Public Use Microdata: Disclosure and Usefulness," Center for Economic Studies Discussion Paper, CES 88-3.

McGuckin, Robert H. and Stephen H. Andrews (1988), "The Performance of Lines of Business Purchased in Conglomerate Acquisitions." Paper was presented at the American Economic Association Meeting in Chicago.

McGuckin, Robert H. and George A. Pascoe, Jr. (1988), "The Longitudinal Research Data Base (LRD): Status and Research Possibilities," Survey of Current Business, Vol. 68, pp. 30-37.
McGuckin, Robert H., Frederick R. Warren-Boulton, and Peter Waldstein (1988), "Analysis of Mergers Using Stock Market Returns," Economic Analysis Group Discussion Paper, EAG 88-1.

Nguyen, Sang V. and Arnold P. Reznek (1989), "Production Technologies, Economies of Scale, and Factor Substitution in Large and Small U.S. Manufacturing Establishments: A Pilot Study," forthcoming in Center for Economic Studies Discussion Paper.

Table 1
Number of Establishments in the LRD for Each Year

YEAR	NUMBER OF ESTABLISHMENTS	NUMBER OF ADMINISTRATIVE RECORD CASES
1963	305,747	*
1967	305,611	118,622
1972	312,398	122,158
1973	73,460	-
1974	68,262	-
1975	71,145	-
1976	70,346	-
1977	350,648	144,648
1978	73,853	-
1979	57,559	-
1980	55,953	-
1981	55,045	-
1982	348,384	128,307
1983	51,619	-
1984	56,551	-
1985	55,128	-
1986	54,858	-

*There were no administrative record cases in 1963.
-There are no administrative record cases in the ASM.

Table 2

Variable in the LRD

Symbol	Variable	Availability*
ppn	permanent plant number	
id	identification number	
ind	tabulated industry code	
ppc	primary product class	
pisr	primary industry specialization ratio	
ppsr	primary product specialization ratio	
il3	status of establishment	
ei	employer identification number	
dind	derived industry code	
et	establishment type (0=ASM)	C
ar	administrative record (1=AR)	C
cc	coverage code	
sc	source code	
lfo	legal form of organization	C
st	state code	
smsa	smsa code	
cou	county code	
plac	place code	
va	value added	
vr	value of resales	
rcw	receipts for contract work	
msc	miscellaneous receipts	
te	total employment	
pw1	production workers: March	
pw2	production workers: May	
pw3	production workers: August	
pw4	production workers: November	
pw	production workers (average)	
ph1	personhours: January-March	
ph2	personhours: April-June	
ph3	personhours: July-September	
ph4	personhours: October-December	
ph	total personhours	
sw	total salaries and wages	
ww	wages: production workers	
ow	wages: other employees	
lc	total supplemental labor costs	

le	legally required supplemental labor costs	
vlc	voluntary supplemental labor costs	
cp	cost of materials, parts, etc.	
cr	cost of resales	
cf	cost of fuels	
ee	cost of purchased electricity	
pe	quantity purchased electricity	
cw	cost of contract work	
cpc	cost of purchased communications	A 77 & 82
fib	b.o.y. inventory: finished goods	
wib	work-in-progress	
mib	materials	
fie	e.o.y. inventory: finished goods	
wie	work-in-progress	
mie	materials	
tib	b.o.y. inventory: total	
tie	e.o.y. inventory: total	
nb	new building expenditures	
nm	new machinery expenditures	
ue	used capital expenditures	
bab	building assets - b.o.y.	A; after 73
mab	machinery assets - b.o.y.	A; after 73
bae	building assets - e.o.y.	A
mae	machinery assets - e.o.y.	A
br	building rents	A
mr	machinery rents	A
bd	building depreciation	A; after 76
md	machinery depreciation	A; after 76
brt	building retirements	A; after 76
mrt	machinery retirements	A; after 76
rbs	building repair	A; after 76
rm	machinery repair	A; after 76
m	material goods	C
mqpc	quantity produced and consumed	C
mqdc	quantity received and consumed	C
mc	delivered cost	C
pi	product code	C
pqp	product quantity produced	C
pqs	product quantity shipped	C
pv	product value shipped	C
pgit	quantity of interplant transfers	C
pvit	value of interplant transfers	C
pqpc	quantity produced and consumed	C
tvs	total value of shipments	C

*The variable is available for all years and all establishments except as noted: A = collected for ASM establishments only;
C = collected in census years only
b.o.y. = beginning of year
e.o.y. = end of year

Table 3

Percentage of Product Class Output in 1981 Switched Out of Product Class in 1982 - Matched Plants*

FREQUENCY OF GROSS CHANGE

VALUE OF SWITCHED OUT OUTPUT, MIDPOINT OF PERCENTAGE CLASS	FREQ	CUM FREQ	PER CENT	CUM PER-CENT
0T***************	110	110	8.23	8.23
1*****************	119	229	8.90	17.13
2****************	110	339	8.23	25.36
3**************	99	438	7.40	32.76
4***************	102	540	7.63	40.39
5*****************************	212	752	15.86	56.25
10******************************	220	972	16.45	72.70
15******************	126	1098	9.42	82.12
20************	83	1181	6.21	88.33
25*******	47	1228	3.52	91.85
30******	36	1264	2.69	94.54
35**	11	1275	0.82	95.36
40***	16	1291	1.20	96.56
45***	14	1305	1.05	97.61
50*	3	1308	0.22	97.83
55**	6	1314	0.45	98.28
60**	5	1319	0.37	98.65
65**	4	1323	0.30	98.95
70**	5	1328	0.37	99.33
75*	1	1329	0.07	99.40
80*	1	1330	0.07	99.48
85*	2	1332	0.15	99.63
90*	1	1333	0.07	99.70
95*	0	1333	0.00	99.70
100R*	4	1337	0.30	100.00

```
   --------+---------+---------+---
          60        120       180

              FREQUENCY
```

*Gross switched out output is calculated by expressing 1981 output of plants producing in the product class that are producing in another product class in 1982 as a percentage of total 1981 product class output.

Table 4

Percentage of Product Class Output in 1981 Switched into Product Class in 1982 - Matched Plants*

FREQUENCY OF GROSS CHANGE

VALUE OF SWITCHED OUT OUTPUT, MIDPOINT OF PERCENTAGE CLASS	FREQ	CUM FREQ	PERCENT	CUM PERCENT
0T*****************	131	131	9.80	9.80
1********************	145	276	10.85	20.64
2*******************	135	411	10.10	30.74
3******************	127	538	9.50	40.24
4***************	95	633	7.11	47.34
5*******************************	226	859	16.90	64.25
10*****************************	209	1068	15.63	79.88
15************	85	1153	6.36	86.24
20**********	64	1217	4.79	91.02
25******	37	1254	2.77	93.79
30***	17	1271	1.27	95.06
35***	17	1288	1.27	96.34
40**	6	1294	0.45	96.78
45**	8	1302	0.60	97.38
50**	7	1309	0.52	97.91
55*	2	1311	0.15	98.06
60*	3	1314	0.22	98.28
65*	2	1316	0.15	98.43
70*	2	1318	0.15	98.58
75*	1	1319	0.07	98.65
80*	0	1319	0.00	98.65
85*	2	1321	0.15	98.80
90*	0	1321	0.00	98.80
95*	3	1324	0.22	99.03
100R**	13	1337	0.97	100.00

```
   --------+---------+---------+----
          60        120       180

              FREQUENCY
```

*Gross switched in output change is calculated by expressing the output of establishments producing in the product class in 1982 and in another product class in 1981 as a percentage of 1981 total output in the product class.

Table 5

Percentage of Product Class Output Due to Net Output Switches in Product Class in 1981 and 1982 - Matched Plants*

FREQUENCY OF NET CHANGE

ABSOLUTE VALUE OF NET OUTPUT CHANGE DUE TO SWITCHES, MIDPOINT OF PERCENTAGE CLASS	FREQ	CUM FREQ	PER- CENT	CUM PER- CENT
0T***************************************	792	792	59.24	59.24
1******	109	901	8.15	67.39
2*****	80	981	5.98	73.37
3****	58	1039	4.34	77.71
4***	49	1088	3.66	81.38
5*****	84	1172	6.28	87.66
10****	65	1237	4.86	92.52
15**	29	1266	2.17	94.69
20**	13	1279	0.97	95.66
25*	8	1287	0.60	96.26
30**	15	1302	1.12	97.38
35*	5	1307	0.37	97.76
40*	4	1311	0.30	98.06
45*	4	1315	0.30	98.35
50*	4	1319	0.30	98.65
55*	0	1319	0.00	98.65
60*	1	1320	0.07	98.73
65*	2	1322	0.15	98.88
70*	1	1323	0.07	98.95
75*	0	1323	0.00	98.95
80*	2	1325	0.15	99.10
85*	1	1326	0.07	99.18
90*	0	1326	0.00	99.18
95*	0	1326	0.00	99.18
100R*	11	1337	0.82	100.00

```
   -----+-----+-----+-----+-----+-----+-----+-----+
       100   200   300   400   500   600   700   800
```

FREQUENCY

*Net output change is calculated by expressing the difference between the output of establishments producing in the product class in 1981 and in another product class in 1982 (switches out) and the output of establishments producing in another product class in 1981 which produced in the product class in 1982 as a percentage of 1981 product class output.

Table 6

Surviving Firms Among Top Firms,
Various Years Survivors in 1982 Among Top Firms in 1977

1977 Industry Concentration Class		Top 20 No.	Top 20 %	Top 8 No.	Top 8 %	Top 4 No.	Top 4 %
Greater than	.6	11.7	58.5	6.0	75.0	3.3	82.5
Between	.4-.6	11.3	56.5	6.0	75.0	3.2	80.0
Less than	.4	9.7	48.5	5.6	70.0	3.0	75.0
Total		11.4	57.0	5.8	72.5	3.1	77.5

Survivors In 1977 Among Top Firms in 1972

1972 Industry Concentration Class		Top 20 No.	Top 20 %	Top 8 No.	Top 8 %	Top 4 No.	Top 4 %
Greater than	.6	12.2	61.0	6.4	80.0	3.5	87.5
Between	.4-.6	11.7	58.5	6.5	81.3	3.5	87.5
Less than	.4	10.4	52.0	5.7	71.3	3.2	80.0
Total		11.8	59.0	6.2	77.5	3.3	82.5

Survivors in 1982 Among Top Firms in 1972

1972 Industry Concentration Class		Top 20 No.	Top 20 %	Top 8 No.	Top 8 %	Top 4 No.	Top 4 %
Greater than	.6	8.8	44.0	5.1	63.8	2.9	72.5
Between	.4-.6	8.0	40.1	5.1	63.8	2.9	72.5
Less than	.4	6.6	33.0	4.1	51.3	2.5	62.5
Total		8.4	42.0	4.7	58.8	2.7	67.5