

NBER WORKING PAPER SERIES

MEASURING THE DYNAMICS OF YOUNG AND SMALL BUSINESSES:  
INTEGRATING THE EMPLOYER AND NONEMPLOYER UNIVERSES

Steven J. Davis  
John Haltiwanger  
Ron S. Jarmin  
C. J. Krizan  
Javier Miranda  
Alfred Nucci  
Kristin Sandusky

Working Paper 13226  
<http://www.nber.org/papers/w13226>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
July 2007

We gratefully acknowledge support from the Kauffman Foundation, the U.S. Census Bureau and the Initiative on Global Financial Markets at the University of Chicago. For helpful comments on earlier drafts, we thank Mark Roberts, Tim Dunne, Thomas Holmes, Robert Strom, Robert Litan, E.J. Reedy, Ying Lowery, Richard Boden, Lucia Foster and participants at the NBER/CRIW conference on Producer Dynamics in April 2005 and an AEA/SBA session in January 2005. We also would like to thank Paul Hanczaryk for helping us understand the Census Bureau's nonemployer data. This work has undergone a more limited review than official Census Bureau publications. The views, findings, and opinions expressed in this work are those of the authors and not the U.S. Census Bureau. All results have been reviewed to ensure confidentiality. The views expressed herein are those of the author(s) and do not necessarily reflect the views of the National Bureau of Economic Research.

© 2007 by Steven J. Davis, John Haltiwanger, Ron S. Jarmin, C. J. Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to

Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes

Steven J. Davis, John Haltiwanger, Ron S. Jarmin, C. J. Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky

NBER Working Paper No. 13226

July 2007

JEL No. C81,D21,L11

**ABSTRACT**

We develop a preliminary version of an Integrated Longitudinal Business Database (ILBD) that combines administrative records and survey data for all employer and nonemployer business units in the United States. Unlike other large-scale business databases, the ILBD tracks business transitions from nonemployer to employer status. This feature of the ILBD opens a new frontier for the study of business formation, early lifecycle dynamics and the precursors to job creation in the U.S. economy. There are 5.4 million nonfarm business firms with employees as of 2000 and another 15.5 million with no employees. Our analysis focuses on 40 industries that account for nearly half of nonemployers and 36 percent of nonemployer revenues. Within these industries, nonemployers account for 14 percent of business revenues. About 220,000 of the seven million nonemployers in our selected industries hire workers and migrate to the employer universe over a three-year horizon. These Migrants account for 20 percent of revenue among young employers (three years or less since first hire). Compared to other nonemployers, the revenue of Migrants grows very rapidly in the year prior to and the year of transition to employer status.

Steven J. Davis  
Graduate School of Business  
The University of Chicago  
5807 South Woodlawn Avenue  
Chicago, IL 60637  
and NBER  
steve.davis@gsb.uchicago.edu

John Haltiwanger  
Department of Economics  
University of Maryland  
College Park, MD 20742  
and NBER  
haltiwan@econ.umd.edu

Ron S. Jarmin  
Bureau of Census  
ron.s.jarmin@census.gov

C. J. Krizan  
Bureau of the Census  
Cornell.J.Krizan@census.gov

Javier Miranda  
Bureau of the Census  
javier.miranda@census.gov

Alfred Nucci  
Bureau of the Census  
Center for Economic Studies  
4600 Silver Hill Road  
Washington, D.C. 20233

Kristin Sandusky  
U.S. Bureau of Census  
LEHD Program  
4700 Silverhill Road  
Suitland, MD 20233  
lee.k.sandusky@census.gov

## **1. Introduction**

The measurement of economic activity by federal statistical agencies focuses greater attention on larger, more mature business units. This data gathering strategy has two clear advantages. First, it yields greater accuracy in estimating the level of economic activity, whether “greater attention” takes the form of higher sampling probabilities or more careful auditing and editing. Second, it is easier to identify and promptly capture the activity of large, long-established business units. On both counts, the desire for a cost-effective approach to measuring the level of economic activity leads naturally to a focus on larger, more mature units.

There are, however, drawbacks to this data gathering strategy. When responses to shocks and new developments in the economy vary systematically with business size or age, a focus on larger and more mature units can yield less accurate, potentially misleading measures of changes in economic activity. As a simple example, consider the situation when younger and smaller business units are relatively sensitive to aggregate shocks. In this case, a cost-effective approach to estimating short-term growth rates can require the over sampling of younger and smaller business units, and there is tension between a sample design optimized for the level of activity and one optimized for the growth rate. More important, the traditional focus on larger and more mature units limits our ability to measure and study the early lifecycle dynamics of businesses and to evaluate theories of business formation, selection and growth.

This paper reports our initial efforts to remedy these drawbacks. We develop a preliminary version of an Integrated Longitudinal Business Database (ILBD) that combines administrative records and survey-based data for all nonfarm employer and nonemployer business units in the United States. In the process, we confront conceptual and practical issues that arise in measuring the importance and dynamic behavior of younger and smaller businesses. We also document some basic facts about younger and smaller businesses. In doing so, we exploit the ability of the ILBD to follow business transitions between employer and nonemployer status. This aspect of the ILBD opens a new frontier for the study of business formation and the precursors to job creation in the U.S. economy.

There are 5.4 million nonfarm business firms with employees in the United States as

of 2000 and another 15.5 million with no employees. Most nonemployer business units are quite small, never become employers, and do not link to the employer universe by way of any ownership relation. Nonemployer businesses account for a modest four percent of aggregate U.S. business revenue within the year, but we show that a substantial share of employers originate as nonemployer businesses.

Our analysis focuses on 40 industries for which smaller and younger businesses play especially important roles. These industries account for nearly half of nonemployer business units in the U.S. economy and 36 percent of nonemployer revenues. Within these industries, nonemployers account for 14 percent of business revenues. In addition, more than 11 percent of employers in these industries are connected by some type of ownership link to the nonemployer business universe within the previous eight years. Many of these linkages reflect nonemployer businesses that expand over time, hire workers, and become employers, but other linkage patterns also arise. For example, some employer and nonemployer business units operate simultaneously under common ownership for many years. In the analysis below, we identify and quantify several types of ownership linkages between the employer and nonemployer business universes.

Over a three-year horizon, three percent of the roughly seven million nonemployers in our selected industries become Migrants to the employer universe. By Migrants, we mean that the firm starts as a nonemployer business and later hires one or more paid employees. Three percent is a small share of nonemployers, but it amounts to 220,000 transitions. Indeed, Migrants play a nontrivial role in the formation of employer businesses. They account for 28 percent of firms and 20 percent of revenues among young employers (three years or less since first hire) in our selected industries. Migrants also grow very rapidly around the transition event. Mean annual revenue growth for Migrants is 31 percent in the year prior to transition and 101 percent in the year of transition, much higher than contemporaneous growth for other nonemployers.

As another step toward an integrated perspective on the dynamics of young and small businesses, we compare the revenue growth patterns of employers and nonemployers. Conditional on survival, net revenue growth rates decline strongly with business age and size for employers and nonemployers alike. When we include business exits, however, revenue growth shows much weaker and less clear-cut relationships to

size and age. The dispersion in growth rates is also much higher for younger and smaller businesses, mainly because of much higher business turnover rates.

The paper proceeds as follows. Section 2 describes the construction of the ILBD. Section 3 presents some facts about business numbers, activity levels, and the business age and size distributions. Section 4 investigates ownership linkages over time between employer and nonemployer businesses. We quantify transitions between employer and nonemployer status and other linkages between employer and nonemployer businesses. Section 5 reports revenue growth and dispersion patterns by business size and age. Section 6 discusses next steps in our research program, and Section 7 offers concluding remarks.

## ***2. Constructing an Integrated Longitudinal Business Database***

### **A. Overview of Main Tasks and Previous Work**

In terms of data development, our objective is to build a fully Integrated Longitudinal Business-level Database (ILBD) that covers all employer and nonemployer business units in the nonfarm private sector of the U.S. economy. We construct the initial version of the ILBD for the years 1992 and 1994-2000, and we plan updates for later years in future ILBD enhancements. Key data on nonemployers are unavailable for 1993.

From an analytical perspective, the presence or absence of employees is simply another business characteristic to be measured. From a database development perspective, however, integrating the Census Bureau's employer and nonemployer business universes is a major undertaking. The main tasks fall into three categories. One set of tasks involve the construction of longitudinal links for business units within each universe. A second task is to integrate the employer and nonemployer universes on a year-by-year basis, ensuring that each unique business entity is counted once, and only once. A third task is to construct contemporaneous and dynamic ownership links across universes between employer and nonemployer business units. To carry out these tasks, we build on previous work by Jarmin and Miranda (2003) to create the Longitudinal Business Database (LBD), which contains annual data from 1975 to 2001 for all nonfarm private employers. We also build on previous efforts to construct longitudinal links

within the nonemployer universe by Nucci and Boden (2003) and Boden and Nucci (2004).

## **B. Source Data for the ILBD**

Census Bureau business registers draw on payroll tax records, corporate and individual income tax returns, applications for an Employer Identification Number (EIN), and various Census Bureau business surveys. The data available to the Census Bureau vary with the legal and tax status of a business and, in certain respects, its size and number of locations. For large corporations, routine data inputs include payroll records and particular items from corporate income tax returns, augmented by direct Census Bureau collections for multi-location companies. For sole proprietors, partnerships and single-location corporations with employees, routine data inputs include payroll records, certain items from income tax returns and periodic Census Bureau surveys such as the quinquennial Economic Census. For nonemployer businesses, routine data inputs derive mainly from income tax returns. Table 1 lists the most important administrative and survey sources for key variables in the employer and nonemployer universes.

To construct the ILBD, we must first ensure that administrative data from each universe are cleaned and ready for integration. On the employer side, this task has been largely accomplished in the work to create the Longitudinal Business Database (LBD).<sup>2</sup> The Census Bureau's Employer Business Register, which underlies the LBD, is a list of establishments (physical locations) maintained to serve as a mailing list for the Economic Census and as a sample frame for surveys. The Employer Business Register relies heavily on administrative data and is supplemented by direct Census Bureau collections.<sup>3</sup> Longitudinal linking is facilitated by establishment IDs (LBD Numbers and Permanent Plant Numbers or PPNs), EINs, enterprise IDs (Alphas), and business name and address information. Longitudinal establishment links are relatively straightforward to construct, because they are one to one, and because establishments typically have well-defined

---

<sup>2</sup>See Jarmin and Miranda (2003). The main outstanding issue with respect to the LBD concerns the delayed identification of new establishments owned by certain multi-unit companies. We are developing algorithms to retime these births. The retiming issue pertains only to the recognition date of establishment birth, not the company-wide level of revenues or other measures of economic activity.

<sup>3</sup>In order to track the establishment structure of multi-unit enterprises, the Census Bureau conducts an annual Company Organization Survey. This survey covers all large multi-unit companies and a sample of smaller ones. During an economic census, all establishments of multi-unit companies receive survey forms.

physical locations. Firms are more difficult to track over time, partly because firm-level links can be many to many. It should be noted that the Census Bureau continues to refine its longitudinal firm linkages.

Longitudinal links are difficult to construct in the nonemployer universe for some of the same reasons that they are difficult to construct for firms in the employer universe. Some nonemployer businesses have an EIN, but most do not and instead are tracked by the person ID of the business owner, i.e., his or her Social Security Number (SSN).<sup>4</sup> When there is a change in the legal or tax status of a nonemployer business, its EIN or person ID can also change. (Person IDs do not change for individuals, but ownership changes can yield a change in the person ID associated with a business.) In these instances, it is not straightforward to maintain longitudinal links for nonemployer businesses using data items that are routinely included in the Census Bureau's administrative records systems. Direct Census Bureau collections provide this additional information on the employer side, but there is no ready equivalent on the nonemployer side. Our longitudinal links for nonemployer businesses currently exploit EINs, person IDs in the form of SSNs, business name information, and geographic information.

Our main source of revenue data for nonemployer businesses are Schedule C forms attached to personal income tax filings. Two complications arise in this regard. First, multiple Schedule C forms can be attached to a single 1040 tax form. In these cases, we aggregate to the level of a single 1040 tax filing and associate the revenue to the SSN of the primary tax return filer. Second, when married couples with Schedule C income file jointly, either the husband or the wife can be the primary filer. In these cases, we associate the Schedule C income to the SSN of the primary filer, but we retain the spouse's SSN as well. We then create longitudinal links based on the SSN of the primary filer. In addition, for married couples who file jointly in year  $t$  and  $t+k$ , we create additional links that capture a primary filer in year  $t$  who matches to his or her spouse's SSN in year  $t+k$  (for  $k$  positive or negative). These additional cases account for less than one percent of all longitudinal linkages among nonemployers.

---

<sup>4</sup> A nonemployer business has an Employer Identification Number if it previously had paid employees or applied for an EIN in anticipation of hiring paid employees.

### **C. Cross-Universe Matching**

In addition to the longitudinal links within each universe, we create a set of firm-level matches between employers and nonemployers for our selected industries. These matches rely on numeric identifiers and exact literal matches on business names. In matching on numeric identifiers, we exploit the fact that many business records contain both an EIN and an SSN. For example, when a business owner or officer applies for an EIN, he or she must fill out an SS-4 form for the IRS. This form includes the business name, the EIN and the SSN of the business owner or chief officer, all of which are included in Census Bureau business registers. These data allow us to build a crosswalk between EINs and SSNs, which we then use to match business records across universes.

We take a conservative approach in matching records between employers and nonemployers. In particular, we rely only on the EIN-SSN crosswalk and exact literal matches on business name. As an example of how our matching algorithm works, consider all establishments with employees in our selected industries as of 2000. Using the longitudinal links in the LBD, we first create a set of identifiers (EINs, SSNs and business names) associated with each establishment with employees in 2000 for each year back to 1992. Given the list of identifiers for a particular establishment, we then identify its matches to nonemployers in the years from 1992 to 2000. Since we are most confident about cross-universe matches based on numeric identifiers, we match on EIN, SSN and business name in that order. For matches on business name, we also require that the employer and nonemployer be in the same state. The industry restriction applies to the business under consideration in the origin universe (i.e., employer universe) and not to its potential matches in the other business universe. We create matches in the other direction using a similar method. That is, for each nonemployer, we match its numeric and name identifiers to the identifiers of establishments in the employer universe.

About 17 percent of our employer-nonemployer matches rely on exact literal matches on business name strings. We experimented with alternative name-matching procedures, including the removal of vowels, various symbols, and commonly used abbreviations (“inc.,” “ltd”, etc.). However, after examining a sample of matches, we concluded that literal name strings produced highly reliable matches and appeared to miss very few “good” matches. Nevertheless, previous efforts to develop longitudinal



identifiers for the LBD have shown that linkages can be improved by the use of more sophisticated probabilistic matching algorithms that exploit all relevant available information, and that take into account the reliability of the information. We plan to incorporate some of these techniques in future enhancements of the ILBD. Given our current algorithm, we probably understate the incidence of linkages between the two universes. For the same reason, we probably understate the contribution of nonemployer businesses to the formation of employer businesses.

Given a set of matches between the two universes, we aggregate the establishment data within an industry to the firm level. The result is a firm-level dataset with ownership linkages to nonemployer businesses and additional variables that describe the nature of the nonemployer records to which the employer firm links. At this point, the unit of observation is a business firm with at least one establishment operating in one of our selected industries. If a firm operates in more than one of our selected industries, we maintain separate records for each industry in which it operates.

### **3. Basic Facts about Employers and Nonemployers**

#### **A. Business Numbers and Activity Levels**

Table 2 provides summary statistics for the employer and nonemployer business universes in 2000. There are about 15.5 million nonemployer businesses. Of these, 13.4 million are person ID units (sole proprietorships with no employees) and 2.1 million are EIN units (corporations, partnerships and other nonemployer business entities with EINs).<sup>5</sup> There are also about 5.4 million employer businesses. Of these, 182 thousand are multi-unit enterprises with more than one establishment, and the rest are single-unit businesses. While comparatively small in number, multi-unit enterprises account for 61 percent of aggregate U.S. business revenue. Nonemployer business units account for four percent of aggregate revenue, and single-unit employers account for 35 percent.

Given the sheer size of the Census Bureau business registers and some complex issues of measurement, we focus on a selected set of 40 industries for this paper. We

---

<sup>5</sup> The distinction between person ID and EIN units can be complex. A sole proprietor with no payroll but positive receipts who has applied for an EIN can appear in both the person ID and EIN sections of the Nonemployer Business Register. That same proprietor can appear in the Employer Business Register as well. We assign all zero-payroll units to the nonemployer universe, even if they reside in the Employer Business Register.

choose industries with large numbers and relatively high revenue shares for young and small businesses. Dynamic links between employers and nonemployers are likely to be more important for these industries. We avoid industries with complex measurement issues related to financial holding companies, tax shelters, and special purpose financial entities. These aspects of corporate organizational structure are interesting, but they are not the focus of our efforts.

Our analysis period overlaps with the transition from SIC to NAICS industry classifications, and the SIC-NAICS crosswalk is a many-to-many mapping. Our nonemployer data files contain 3-digit SIC codes prior to 1997 and 4-digit NAICS codes thereafter. The employer data files contain codes for both classifications from 1997 to 2000. Accordingly, we proceed as follows. For many exercises, we look backwards for businesses in selected 4-digit NAICS industries. For other exercises, we look forward from a year prior to the NAICS changeover at businesses in 3-digit SIC codes that correspond closely to our selected NAICS industries.

Table 3 provides summary information for our selected 4-digit NAICS industries. Legal Services has the largest number of employer businesses, almost 150 thousand. It also has the biggest employment and labor costs, with more than one million workers and more than 58 billion dollars in payroll. The highest-revenue industry for employers is Gasoline Stations at 187 billion dollars.<sup>6</sup> Other Personal Services has the largest number of nonemployer businesses, more than 800 thousand. The highest-revenue industry for nonemployers is Real Estate Agents and Brokers at almost 23 billion dollars.

Table 4 provides information about industry shares of aggregate business revenues and the relative size of the employer and nonemployer segments within industries. Nonemployer revenue shares range widely. At the upper end, nonemployers account for more than two-thirds of revenue in Independent Artists, Writers and Performers, and they account for at least 30 percent of revenues in a dozen industries. At the lower end, nonemployers account for a mere 24 hundredths of a percent of revenue in Software Publishers, even though 20 percent of all businesses in this industry are nonemployers. Although extreme, the basic pattern in this industry is not uncommon; it reflects the

---

<sup>6</sup> High revenue in Gasoline Stations mainly reflects the cost of gasoline. The administrative data in the Census business registers typically does not include information on gross margins or material costs. Such data are included in the Economic Censuses and various annual surveys.

enormously skewed size distribution of activity in many industries.

## **B. Business Age and Size Distributions**

Figures 1 to 4 provide information about the age and size distributions of businesses in our selected industries. Age is measured in years since first appearance by a business entity in its respective business universe. For multi-unit firms, business age is defined as the age of the firm's oldest establishment. We use revenue measures to compare size distributions across the two universes, because revenue is the only activity measure available for both universes.

As seen in Figure 1, older firms dominate economic activity among employers. Firms that are at least eight years old account for almost 70 percent of employer revenues and more than 40 percent of employer businesses in our selected industries. In contrast, older businesses play a much smaller role in the nonemployer universe. Firms that are at least eight years old account for only 40 percent of nonemployer revenues. Very young nonemployers account for a larger share of business units and even revenue than, say, five-year old nonemployers.

As seen in Figure 2, most nonemployer businesses are quite small. Roughly a third of nonemployer businesses generate less than 6,000 dollars in annual revenue. While large in numbers, these very small nonemployers account for a tiny fraction of revenues. Somewhat larger business units account for much of the revenue generated by nonemployers. For example, nearly one-quarter of nonemployer revenue in our selected industries is generated by businesses with annual revenue in the range of 120 to 360 thousand dollars. In contrast, the size distribution of revenues has a very different shape in the employer universe. Almost 70 percent of employer revenue is generated by firms with more than three million in annual revenue.

The tremendous variation in size across nonemployer and employer businesses exhibited in Figure 2 serves as a caution when drawing inferences about the behavior of "small and young" businesses. The wide size distribution reminds us that many nonemployer businesses are extremely small and often represent a secondary or supplemental source of income to the household. Analyzing the dynamics of such businesses alongside much larger businesses is a challenge. In what follows, we often

report results for both the share of business units and the share of revenue. The former provides more insights about the very small and more prevalent businesses, while the latter provides more insights into the contribution of larger businesses.

Figures 3 and 4 display the share of revenues and business units accounted for by young businesses (0-3 years old) and small businesses (less than \$90,000 in annual revenue) in our selected industries. Figure 3 reveals wide variation across industries in the revenue and number shares of young and small businesses in the nonemployer universe. Figure 4 shows a similar pattern with respect to the revenue and numbers share of young businesses in the employer universe and with respect to the numbers share of small businesses. However, with the exception of Personal Care Services (NAICS 8121), employers with less than \$90,000 in annual revenue account for very small revenue shares, typically less than five percent.

#### ***4. Ownership Links and Transition Dynamics***

##### **A. Backward Links of Employers to Nonemployers**

We turn now to ownership links between the employer and nonemployer business universes. We first consider all employer businesses in our selected industries in the 2000 cross section. Table 5 reports size and age distributions for these businesses in columns (2) and (3). These columns show familiar patterns: the number of active businesses declines with age and size, but the bulk of activity – here measured by payroll – is concentrated in older and larger businesses.

The more novel elements of Table 5 appear in the last four columns, which provide information about ownership links between the employer and nonemployer universes. The at-risk population for this analysis is all employer businesses in our selected industries with positive payroll in 2000. For these employers, we consider all ownership links to nonemployers in the current and previous eight years, i.e., in 1992 or 1994 to 2000. Out of the 2.3 million employers in our selected industries, about 266 thousand have ownership links to the nonemployer universe within the current or previous eight years based on the matching algorithm described above.

Columns (4) and (5) in Table 5 report the number and percentage of employer firms

with ownership links to nonemployers by employer size and age. While most employers have no ownership links to the nonemployer universe, many do, and this pattern holds for all size and age categories. Among firms with 1 to 4 employees, 14 percent link to the nonemployer universe within the previous eight years. The propensity for links to the nonemployer universe is U-shaped in employer size. Among firms less than six years old, more than 15 percent link to the nonemployer universe. The propensity for ownership links to nonemployer businesses declines after age five, but this pattern may simply reflect our inability to identify ownership links in 1993 and prior to 1992.

Column (7) in Table 5 reports the pre-link distribution of nonemployer revenues for those nonemployers that link to the employer universe. To construct this distribution, we sum deflated revenues over the nonemployer records that link to the 2000 LBD. We use the nonemployer's revenue value in the year prior to the link. For example, if a 1997 nonemployer record links to an employer that operates in 2000, we use the deflated 1996 nonemployer revenue in the computation. Since there can be multiple dynamic links between employer and nonemployer, we take the oldest link and count each nonemployer at most once. Column (6) is constructed in the same manner as column (7), except that each nonemployer record receives a unit weight. Comparing columns (6) and (7) in the top panel yields the inference that relatively large nonemployers tend to link to larger employers. To see this, note that employers with at least 500 workers account for 6.2 percent of pre-link revenues among linked nonemployers but only 0.24 percent of the linked nonemployers. Similarly, over 75 percent of the employer businesses with links to a nonemployer business have fewer than five employees, but they link to less than 48 percent of pre-link nonemployer revenues.

When thinking about the process of business formation and growth, we anticipate a pattern whereby some businesses start as nonemployers, grow over time, and eventually transition to employer status, perhaps continuing to grow thereafter. This pattern holds for many businesses in the ILBD, but it is certainly not the only linkage pattern that arises. This point is evident in the lower panel of Table 5, specifically in columns (6) and (7). More than 30 percent of nonemployer firms that link to the employer universe – and more than 45 percent of pre-link revenues – involve links to employers that are at least eight years old as of 2000. All of these cases involve nonemployer firms that link to previously established employer businesses. That is, they do not involve a nonemployer

business that evolves into a new employer business.

Figure 5 highlights this point in greater detail by plotting the frequency distribution of the age difference between employers in the 2000 cross section and their linked nonemployers (again using the first link to the employer). Recall that we know the exact age in years for employers that first appear in the employer universe after 1975. For nonemployers, we construct an age measure based on first appearance in the Nonemployer Business Register. Since we only begin observing nonemployers in 1992, examining this age difference for nonemployers observed first in 1992 is clouded by left censoring issues. As such, for Figure 5, we examine all 2000 employers with links to nonemployers where the nonemployer first appeared after 1992 so that we have an accurate age for the nonemployer.<sup>7</sup> For each of these linked nonemployers, we then compute the difference between its age and the age of the employer to which it links.

According to Figure 5, 60 percent of nonemployers are older than the employer to which they link. These cases are consistent with the “standard” pattern whereby a nonemployer business evolves into a new employer business. The pronounced mode at a one-year age difference reflects businesses that transition to employer status one year after inception as a nonemployer. Many other businesses operate in nonemployer mode for a few years before transitioning to employer status. In addition to these standard cases, Figure 5 shows a large number of linkages in which the employer business pre-dates the nonemployer business. These “nonstandard” linkage cases reflect other types of ownership relations between the two business universes. For example, an individual who owns a business with employees may also generate consulting income in a nonemployer business. As another type of example, a corporate business with employees may establish nonemployer subsidiaries for legal, financial or tax reasons.

## **B. Nonemployer Transitions**

To continue our exploration of linkages between the two business universes, we now conduct an analysis of transitions. We first examine transitions from the nonemployer universe. In particular, we consider the population of 1994 nonemployer businesses in our selected industries and classify their operational status three years later in 1997.

---

<sup>7</sup> We have also examined a version of Figure 5 with the left censored cases included and the results are very similar to those reported.

Figure 6 summarizes the three-year transition dynamics for the population of nonemployer businesses in one of our selected industries.<sup>8</sup> The at-risk population is all nonemployer businesses with revenues in 1994. We classify outcomes into six categories:

- **Exits** – businesses with positive revenue in the nonemployer universe in 1994, no revenue in the nonemployer universe in 1997, and no payroll in the employer universe in 1997.
- **Migrants** – businesses with positive revenue in the nonemployer universe in 1994, positive payroll in the employer universe in 1997, no revenue in the nonemployer universe by 1998,<sup>9</sup> and the matched employer does not pre-date the nonemployer with respect to the age of its oldest establishment.
- **Other Transits** – businesses with positive revenue in the nonemployer universe in 1994, positive payroll in the employer universe in 1997, no revenue in the nonemployer universe by 1998, and the matched employer pre-dates the nonemployer business.
- **Continuers** – businesses with positive revenue in the nonemployer universe in 1994 and 1997, and no payroll in the employer universe in 1997.
- **Persistent Duals** – businesses with positive revenue in the nonemployer universe in 1994, 1997, and 1998 and positive employer payroll in 1997 and 1998.
- **Other Duals** – business with positive revenue in the nonemployer universe in 1994 and 1997, positive payroll in the employer universe in 1997, no revenue in the nonemployer universe in 1998 and no payroll in the employer universe in 1998.

We compute the share of the 1994 nonemployer analysis population that falls into each category on an unweighted and revenue-weighted basis.

Continuers account for 62 percent of nonemployer businesses and 58 percent of their

---

<sup>8</sup> The basic patterns for these transition dynamics are very similar over a six-year horizon, although the magnitudes change in the expected way, e.g., the share of activity accounted for by exits rises substantially.

<sup>9</sup> We require no revenue in the nonemployer universe by 1998, rather than 1997, because a business that transitions during the 1997 calendar year will typically record positive 1997 revenues in both universes.

revenues in 1994. Exits account for another 38 percent of nonemployers and 26 percent of their revenues. Migrants account for only three percent of nonemployer businesses but nine percent of their 1994 revenues. While three percent is a small share of the population at risk, there are 7.4 million nonemployer businesses in our selected industries. In terms of raw numbers, approximately 220,000 nonemployers in 1994 migrate to employer status by 1997. Other Transits, Persistent Duals and Other Duals account for very small shares of businesses and revenue. Figure 7 shows that Migrants, and Exits occur with greater frequency among EIN cases than SSN cases.

### **C. The Contribution of Migrants to Young Employers**

We now quantify the contribution of Migrants to the economic activity of young employers with positive revenue in 1997 but no paid employees prior to 1995. Thus young employers have had paid employees for at most three years. As reported in Table 6, Migrants account for 28 percent of young employers and 20 percent of their revenues in our selected industries. Several industries stand out for the large role of Migrants including Farm Labor and Management Services, Landscape and Horticultural Services, Painting and Paper Hanging, Carpentry and Floor Work, Insurance Agents and Brokers, Real Estate Agents and Managers, Computer and Data Processing Services, Automotive Repair Shops, Legal Services, Child Day Care Services, and Accounting, Auditing and Bookkeeping. Migration from nonemployer to employer status is an important feature of business formation and growth in these industries.

### **D. The Growth Pattern of Migrants**

Next, we compare the growth rates of Migrants to two other groups of nonemployers – Continuers plus Exits and Continuers Only. For this purpose, we restrict attention to Migrants with nonemployer status as of 1996 that transit to employer status in 1997. That is, we consider Migrants with positive revenue in the nonemployer universe in 1996, positive revenue in the employer universe in 1997, and no revenue in the nonemployer universe by 1998. As before, we also require that a Migrant pre-date its matched employer. For Migrants with positive activity in both universes in 1997, its 1997 revenue value is the sum of its nonemployer and employer revenues. We also limit Exits and Continuers in the comparison groups to those with positive revenue in the 1996



nonemployer universe. We measure the growth rate as the change in annual revenue from  $t-1$  to  $t$  divided by the simple average of revenue in  $t-1$  and  $t$ .<sup>10</sup> We compute all summary statistics on a revenue-weighted basis.

As reported in Table 7, Migrants grow much faster than other nonemployers. In the year prior to transition, the mean (median) growth rate for Migrants is 31 (15) percent, compared to 20 (5) percent for Continuers. In the year of transition, the mean (median) growth rate for Migrants is 101 (102) percent, compared to 6 (3) percent for continuers. In short, Migrants are on a trajectory of rapid growth before and during the transition to employer status.

## **E. Employer Transitions**

Figure 8 summarizes three-year transition dynamics for the 1994 population of employers in our selected industries. As before, we group businesses in the at-risk population into six categories based on their status three years later. The categories mirror the ones considered previously. For example, “Migrants” now refer to businesses with positive revenue in the employer universe in 1994, positive revenue in the nonemployer universe in 1997, no revenue in the employer universe by 1998, and the employer pre-dates its matched nonemployer.

Continuers – businesses with positive revenue in the employer universe in both years and no ownership links to the nonemployer universe in 1997 – account for 68 percent of all employers and 74 percent of their 1994 revenues. Exits account for 19 percent of employer revenues and 27 percent of employer businesses. The exit figures point to high death rates for employers in our selected industries, but they are considerably smaller than exit rates for nonemployers (Figure 6). There are approximately 39,000 Migrants from the 1994 employer universe to the 1997 nonemployer universe, which amounts to about two percent of employers and two percent of their 1994 revenues. Other Transits account for one percent of employer businesses and revenues. Persistent Duals account for one percent of employer businesses and three percent of employer revenues.

---

<sup>10</sup> The growth rate measure is bounded, symmetric about zero and ranges from -200 to 200 percent, with endpoints corresponding to exit and entry. See Davis, Haltiwanger and Schuh (1996).

## **5. Revenue Growth and Dispersion by Age and Size**

There is a vast body of research on the relationship of business growth patterns to business size and age. Almost all of this research restricts attention to businesses with employees, and much of it considers a subset of employers that meet a minimum size threshold or that include only publicly traded companies.<sup>11</sup> Using the ILBD, we can systematically analyze and compare the dynamics of employer and nonemployer businesses.

We consider business growth from 1999 to 2000 with attention to mean growth rates by size and age and the dispersion of growth rates within size and age categories. Our dispersion measure is the excess revenue reallocation rate: gross revenue gains at expanding units plus gross revenue losses at contracting units minus the absolute value of the net revenue change, all divided by aggregate revenue for the units under consideration. The excess reallocation rate is equivalent to the average absolute deviation of growth rates about zero, confirming its interpretation as a measure of cross-sectional dispersion in growth rates.<sup>12</sup> We consider all employers in our selected industries but limit attention to Continuers, Exits and Migrants for the nonemployers. These three groups account for the vast majority of nonemployers and their revenues (Figures 6 and 7). For Migrants from nonemployer to employer status, our 2000 revenue measure includes nonemployer revenue, if any. In principle, we could treat Migrants from employer to nonemployer status symmetrically, but we ignore the matter as unimportant.

Figure 9 shows that the mean growth rate of employers drops off very rapidly by age two and displays no clear relationship to age among older employers when we include Exits. The drop in mean growth with age is even more rapid among nonemployers. Indeed, mean nonemployer growth rates are negative for all ages beyond zero, i.e., beyond the year of entry. Conditional on survival, mean growth is positive at all ages and shows a clear tendency to decline with age.

Figure 10 shows that excess revenue reallocation rates tend to decline with business age, especially for nonemployers. Perhaps more important, the magnitude of excess

---

<sup>11</sup> Dunne, Roberts and Samuelson (1989), Sutton (1997), Caves (1998) and Davis and Haltiwanger (1999) review various branches of this literature, which spans several decades.

<sup>12</sup> See Davis and Haltiwanger (1999), who review the use of this measure in the literature on job flows.

revenue reallocation is very large for employers and nonemployers alike, and at all ages. Excess revenue reallocation exceeds 50 percent in all age groups for nonemployers. It exceeds 30 percent in all age groups for employers. These results underscore the tremendous amount of revenue expansion and contraction that takes place on a routine basis among U.S. businesses. In this respect, the results echo previous findings on the large magnitude of simultaneous job creation and destruction in Davis, Haltiwanger and Schuh (1996) and other work. When we condition on survival, considerable revenue churning remains, but it does not show a strong relationship to business age.

Figures 11 and 12 display revenue growth and excess reallocation rates by business size. The size categories are narrow at the lower end to reflect the revenue distribution among nonemployers. As seen in Figure 11, mean growth rate for small revenue classes is highly sensitive to whether we restrict attention to survivors. Conditional on survival, very small businesses have very high net growth rates relative to their larger counterparts. When we include Exits, the relationship between net revenue growth and size is basically flat.

Figure 12 shows that excess revenue reallocation rates decline sharply with size for employers and nonemployers. Excess reallocation is high for businesses of all sizes, exceeding 20 percent even for the largest businesses. Perhaps surprisingly, excess reallocation rates among businesses with less than \$120,000 in annual revenue are greater for employers than nonemployers. However, we know from Figure 2 that there is little revenue activity in the very small size classes. Once again, conditioning on survival has a profound effect on size the relationship, although excess revenue reallocation remains high in all size categories.

## **6. *Where Do We Go From Here?***

The preceding sections describe the employer and nonemployer business universes, relate our efforts thus far to integrate the two universes, and present evidence on the dynamics of young and small businesses including the migration of nonemployers to employer status. In this section, we describe several challenges that arise in further developing the ILBD.

## **A. Issues Common to Both Business Universes**

One issue common to both business universes is the conversion from SIC to NAICS industry codes following the 1997 economic censuses. A related but larger set of problems involves the reliability and accuracy of industry codes in the ILBD. Large portions of the two business universes rely almost exclusively on administrative records for source data on industry codes. As a rough generalization, the industry codes are less reliable and less precise for nonemployers and for smaller employers. Geographic identifiers also tend to be less accurate for nonemployers and smaller employers. In general, there are fewer sources of information for business-level records that derive entirely from administrative sources, as compared to those that rely on administrative and survey sources.

Another common issue pertains to the interpretation of business revenue measures and their consistency over time. In particular, revenue measures can be affected by changes over time in income tax rules. This issue merits study.

## **B. Issues in the Employer Universe**

Multi-unit businesses above a size threshold are surveyed by the Census Bureau in the annual Company Organization Survey (COS). However, the list of such businesses is drawn from the prior economic census. These procedures mean that a firm's transition from single-unit to multi-unit status often goes undetected until the next economic census. In addition, new establishments operated by small multi-unit firms not covered by the COS are detected only at the economic censuses. In both cases, the economic activity measures for these new establishments are included with older establishments of the parent company in the inter-censal years. Hence, the delayed recognition of some new establishments in inter-censal years leads to inaccurate establishment counts and, possibly, to an initially incorrect geographic and industrial classification for these new establishments. These issues are not critical for this paper because our unit of analysis is the firm, but they are important for the development of the LBD and ILBD.

Turning to another issue, the Census Bureau has made considerable progress in developing and maintaining longitudinal establishment identifiers for employer businesses, but the development of firm-level longitudinal identifiers remains an open

area for research and development. Standard firm-level identifiers automatically change when a business undergoes certain types of reorganization such as a change in its legal form of organization or a merger. In the analysis above, we dealt with this issue by equating firm age to the age of the oldest establishment operated by the firm. However, we identified entry and exit of firms based on changes in standard firm-level identifiers in the ILBD.

### **C. Issues in the Nonemployer Universe**

Some data issues unique to the nonemployer universe reflect the relatively recent availability and development of annual nonemployer files at the Census Bureau. For example, we cannot trace the inception of nonemployer businesses to years before 1992. In addition, as mentioned above, the nonemployer data rely very heavily on administrative sources. The construction of longitudinal links for nonemployer business units also raises several challenges, and our work in this area is at a relatively early stage of development. There is room for improving the longitudinal and cross-sectional linkages via name and address matching, the treatment of joint returns for proprietorships (where there are separate firm identifiers for the filer and his or her spouse), and the reliability of employer identification information for nonemployer proprietors.

### **D. Integrating the Two Business Universes**

Studies of the integrated business-level data also face other challenges. First, the standards for classification by industry and geography differ between the two universes. These differences limit our ability to isolate narrowly defined industries and regions. Second, at the most basic level, the unit of observation differs between the two universes. For employers, the fundamental unit of observation is typically an establishment. For nonemployers, it is a tax return that reflects economic activity at the home or other locations. Our current analysis also aggregates tax filers with multiple Schedule C forms into a single nonemployer entity, even when each Schedule C involves quite different business activities. Third, some firms with employees create affiliated business entities with no employees in order to shelter income from taxation or limit legal exposures and financial risks. We deliberately sought to sidestep the complex measurement issues associated with these special-purpose business entities through our choice of industries.

A thorough treatment of this issue for all industries is likely to require careful study of the legal framework and economic incentives governing the creation of special-purpose business entities.

## **E. Integrating Employee Records with the ILBD**

An exciting direction for future research is the integration of *employee* data with the ILBD. Using the longitudinal matched employer-employee data from the LEHD program at the Census Bureau, demographic and earnings data for the universe of employees can be integrated with the ILBD files.<sup>13</sup> Integration of the employer, nonemployer, and employee data will provide an unprecedented opportunity to study firm, job and worker dynamics. For example, it will be possible to follow someone who first works as an employee in a specific industry, then starts a small nonemployer business on the side, and later opens an employer business. More generally, for questions about where and how employer businesses originate, it will be very useful to know the demographic characteristics of business founders and their previous histories as employees and business owners.

## **7. Concluding Remarks**

It is tempting to think of the nonemployer business universe as a vast nursery for employer businesses. According to this view, many nonemployers evolve into employers and a few eventually grow into giant corporations that generate thousand of jobs. However, as our results confirm, most nonemployer business are quite small and never become employers. Indeed, it is misleading to think of all records in the nonemployer universe as “businesses” in the usual sense. Many nonemployer records reflect side jobs, hobby businesses or occasional consulting engagements that generate extra income for households that depend primarily on wages. One goal of our research is to help the Census Bureau develop algorithms that can distinguish hobby businesses, for example, from other types of nonemployer businesses, including entrepreneurial undertakings that might evolve into larger businesses with employees.

---

<sup>13</sup> Another important direction for future work is the integration of the Characteristics of Business Owners (CBO) and Survey of Business Owners (SBO) data sets into the ILBD. Holmes and Schmitz (1995), amongst others, have shown the rich analysis that can be conducted with the CBO.

There is, however, a kernel of truth in the “vast nursery” view. Our evidence shows that migrants from the nonemployer universe account for a sizable share of young employers in the industries we study. These Migrants make up 28 percent of young employers (zero to three years old) and account for 20 percent of their revenues. Their importance varies considerably across industries. Among young employers, Migrants account for 38 of revenues in Real Estate Agents and Managers, 35 percent in Taxicabs, and 30 percent in Painting and Paper Hanging and Landscaping but only 11 percent in Eating and Drinking Places. These figures probably understate the role of Migrants because of our conservative matching algorithms. In any event, the results indicate that a significant fraction of employers originate as nonemployer businesses.

On the data front, this study takes important strides in developing an Integrated Longitudinal Business-level Database. Considerable work lies ahead, but the ILBD is already yielding useful information about the dynamics of young and small businesses. Major strengths of the ILBD include comprehensive industry and geographic coverage, longitudinal links for establishments and firms, linkability to the large number of business surveys housed at Census and an integrated treatment of employer and nonemployer business. The ILBD makes it possible to examine the behavior over time of virtually all businesses in the U.S. economy, employers and nonemployers alike, with robust samples and even entire populations.

## References

- Boden, Richard and Alfred Nucci, "Business Dynamics among the Smallest and Newest of Businesses: Preliminary Analyses of Longitudinally Matched Nonemployer and Business Register Data." Annual Meetings, Eastern Economic Association, Washington, D.C. March, 2004.
- Caves, Richard E., "Industrial Organization and New Findings on the Turnover and Mobility of Firms" *Journal of Economic Literature*, 36 (4), 1998.
- Davis, Steven J. and John Haltiwanger, "Gross Job Flows" in *Handbook of Labor Economics* (Orley Ashenfelter and David Card, eds.), North Holland, 1999.
- Davis, Steven J., John C. Haltiwanger, and Scott Schuh, *Job Creation and Destruction*, MIT Press, 1996.
- Dunne, Timothy, Mark Roberts and Larry Samuelson, "The Growth and Failure of U.S. Manufacturing Plants" *Quarterly Journal of Economics*, November 1989.
- Holmes, Thomas and James Schmitz, "On the Turnover of Business Owners and Business Managers," *Journal of Political Economy*, October 1995.
- Jarmin, Ron, and Javier Miranda, "The Longitudinal Business Database," CES Working Paper, 2003.
- Nucci, Alfred and Richard Boden, "Demography of Nonemployer Businesses -- Preliminary Evidence from the United States." Paper Presented at the Comparative Analysis of Enterprise (micro) Data Conference, London, England, September 15-16, 2003
- Sutton, John, "Gibrat's Legacy," *Journal of Economic Literature*, 35 (1), 1997.



**Table 1.A: Employer Business Register**

<b>Variable</b>	<b>Source Type</b>	<b>Source Details</b>	<b>Line Number</b>
<b>Name and Address</b>	Survey Administrative	Physical address from Company Organization Survey, Annual Surveys, or Economic Censuses in Census years Beginning in 1998, physical address from form ss-4 for births; before 1998, mailing address from form ss-4 Beginning in 1998, physical address from form 941 for all businesses; before 1998, mailing address from form 941 for all businesses Beginning in 1998, physical address from IRS income tax form for all businesses; before 1998, mailing address from IRS income tax form for all businesses	4a,4b; 5a,5b Top of form Top of form
<b>Industry</b>	Survey  Administrative	Reported industry code from Company Organization Survey, Annual Surveys, or Economic Censuses in Census years Derived from the 1992 Economic Census--respondent reported classification Derived from a current survey (County Business Patterns (CBP), Company Organization Survey (COS)/Annual Survey of Manufactures (ASM), Current Industrial Report (CIR), Business Sample Revision (BSR) CBP Safeguard Review or intercensal refiles Derived from the Bureau of Labor Statistics Derived from the Social Security Administration Derived from the Internal Revenue Service Principal Business Activity code, obtained from the Business Master File	
<b>Employment</b>	Survey* Administrative	Company Organization Survey, Annual Surveys, or Economic Censuses in Census years IRS Form 941 IRS FICA wages IRS total compensation Imputed	Line 1
<b>Payroll</b>	Survey* Administrative	Company Organization Survey, Annual Surveys, or Economic Censuses in Census years IRS Form 941 IRS FICA wages IRS total compensation Imputed	Line 2
<b>Revenue</b>	Survey* Imputed from EIN-level data	Company Organization Survey, Annual Surveys, or Economic Censuses in Census years  1120 – Gross receipts or sales less returns and allowances 1120-A - Gross receipts or sales less returns and allowances  1120F - Gross receipts or sales less returns and allowances 1120L - Gross income 1120-PC - Gross income 1120-RIC - Total income 1120S - Gross receipts or sales less returns and allowances 1065 – Gross receipts or sales less returns and allowances 990 - Total revenue 990-C - Gross receipts or sales less returns and allowances 990EZ - Total revenue 990-PF - Total revenue 1040C - Gross receipts or sales less returns and allowances	Line 1c Line 1c Section II, line 1a Line 9 Sch A: line 14 Line 8 Line 1c Line 1c Line 9 Line 12 Line 1c Line 9 Line 12 Line 3

\*Data for multi-unit firms disaggregated to the establishment level using survey information, except where the firm reports individual establishments under separate EINs.

**Table 1.B: Nonemployer Business Register, Proprietors - SSN Records**

<b>Variable</b>	<b>Source from IRS Form 1040 including Schedule C</b>
<b>Name</b>	Form 1040
<b>Mailing Address</b>	Form 1040
<b>Legal Form of Organization</b>	Implied by filing of IRS form 1040 Schedule C
<b>Industry Code</b>	Line B
<b>Revenue</b>	Line 3: Gross receipts or sales less returns and allowances

Notes:

1. All data for sole proprietors (including proprietorships jointly operated by husband and wife) are obtained from IRS form 1040 including Schedule C.
2. Line B of Schedule C reports the "Principal business or profession, including product or service." Based on this information, the IRS codes the Principal Business Activity (PBA) of the proprietorship. Based on the PBA code, the Census assigns a Tabulated Kind of Business (TKB) code. If the PBA is not reported on the tax form, then the Census uses the historic TKB code, if available.

**Table 1.C: Nonemployer Business Register, Partnerships and Corporations – EIN Records**

<b>Form Number</b>	<b>Form Description</b>	<b>Industry Code Source</b>	<b>Revenue Source</b>
<b>Form 1065</b>	U.S. Partnership Return of Income	Line A (Principal Business Activity)	Line 1c: Gross receipts or sales less returns and allowances
<b>Form 1120</b>	U.S. Corporation Income Tax Return	Schedule K, Line 2a (Business Activity Code)	Line 1c: Gross receipts or sales less returns and allowances
<b>Form 1120A</b>	U.S. Corporation Short Form Tax Return	Part 2, line 1a (Business Activity Code)	Line 1c: Gross receipts or sales less returns and allowances
<b>Form 1120 S</b>	U.S. Income Tax Return for an S Corporation	Line B (Business Code)	Line 1c: Gross receipts or sales less returns and allowances
<b>Form 1120 F</b>	U.S. Income Tax Return of a Foreign Corporation	Line F1 (Business Activity Code)	Section II, 1c: Gross receipts or sales less returns and allowances
<b>Form 1120 PC</b>	U.S. Casualty and Property Insurance Company income Tax Return	Schedule I, Line 2	Schedule A: line 14 (gross income)
<b>Form 1120 L</b>	U.S. Life Insurance Company Income Tax Return	Schedule M, 2a-c (Kind of Company, Principal Business)	9 Gross Income
<b>Form 1120 RIC</b>	U.S. Income Tax Return for Regulated Investment Companies	Inferred by Form Type	8: Total Income
<b>Form 1120 REIT</b>	U.S. Income Tax Return for Real Estate Investment Trusts	Inferred by Form Type	8: Total Income

Notes:

1. All data for these businesses are derived from IRS income tax returns filed by the businesses
2. Name and Mailing Address taken from top of forms.
3. Industry codes on Census Nonemployer database are IRS PBA codes obtained from the sources noted above and then converted to Census TKB codes. If the PBA is not available from the tax form, then the Census use the historic TKB code, if available.
4. Legal form of organization implied by type of form submitted: 1065 filers are partnerships, all others are corporations.

**Table 2: Summary Statistics for the Employer and Nonemployer Business Universes, 2000**

		<b>Nonemployer Business Universe</b>			<b>Employer Business Universe</b>				
		<u>All Industries</u>	<u>40 Selected Industries</u>		<u>All Industries</u>	<u>40 Selected Industries</u>			
			<i>Number or Revenue</i>	<i>As Percent of All</i>		<i>Number or Revenue</i>	<i>As Percent of All</i>		
<b># of Units (millions)</b>	SSN Units	13.38	6.84	51	Single-Unit	5.26	1.9	36	
	EIN Units	2.15	0.54	25	Multi-Unit	0.18	0.06	31	
	All	15.54	7.38	48	All	5.44	1.96	36	
<b>Revenue (billions \$)</b>	SSN Units	459.53	199.87	43	Single-Unit	6,113.43	877.92	14	
	EIN Units	251.74	55.58	22	Multi-Unit	10,758.04	664.52	6	
	All	711.26	255.45	36	All	16,871.47	1,542.44	9	
<b>Integrated Business Universe</b>									
		<u>All Industries</u>			<u>40 Selected Industries</u>				
		<i>Nonemployer Businesses</i>		<i>Employer Businesses</i>	<i>Nonemployer Businesses</i>		<i>Employer Businesses</i>		
		<u>EIN Units</u>	<u>SSN Units</u>	<u>Single-Unit</u>	<u>Multi-Unit</u>	<u>EIN Units</u>	<u>SSN Units</u>	<u>Single-Unit</u>	<u>Multi-Unit</u>
<b>Percent of Aggregate Revenue</b>		1.43	2.61	34.77	61.19	3.09	11.12	48.83	36.96

**Table 3: Summary Statistics for Selected Industries in the Integrated Business Universe, 2000**

Industry Description	Employers				Nonemployers	
	Firms	Payroll	Employment	Revenue	Firms	Revenue
Animal production support activities	1.5	156	7	664	37.7	1,462
Painting & wall covering contractors	36.4	6,274	226	17,469	213.5	7,443
Carpentry & floor contractors	52.3	9,637	383	35,398	389.6	16,722
Roofing, siding, & sheet metal contractors	27.6	8,143	264	27,315	86.5	5,047
Concrete contractors	24.5	5,092	178	18,185	42.6	2,567
Printing & related support activities	23.2	14,226	420	41,613	26.7	1,486
Ship & boat building	1.1	3,788	105	16,896	0.4	47
Gasoline stations	52.2	12,282	837	187,841	9.4	1,682
Book, periodical & music stores	9.4	2,713	222	12,577	28.9	1,008
Florists	20.7	1,645	123	6,417	22.7	869
Taxi & limousine service	5.1	1,206	65	3,451	117.6	3,419
Couriers	2.3	15,654	550	18,610	1.1	111
Software publishers	6.7	23,009	249	49,988	1.7	120
Agencies & other insurance related activities	101.7	30,448	745	90,461	308.2	14,849
Offices of real estate agents & brokers	50.0	11,504	297	46,826	476.6	22,952
Activities related to real estate	39.5	15,052	484	38,058	356.3	18,274
Consumer goods rental	13.7	3,483	236	11,391	16.0	768
Legal services	147.7	58,514	1,055	149,400	206.3	11,626
Accounting, tax prep, bookkeep, payroll service	76.9	21,273	765	42,205	294.5	6,059
Computer systems design & related services	74.9	76,674	1,194	141,900	249.4	9,688
Management, sci & tech consulting services	76.5	32,718	729	67,277	355.2	16,796
Travel arrangement & reservation services	21.3	8,713	292	45,546	31.8	1,776
Services to buildings & dwellings	115.2	24,903	1,407	59,490	538.9	11,294
Offices of physicians	120.5	46,346	1,043	102,651	149.6	11,664
Offices of dentists	35.0	7,268	273	19,865	29.4	1,907
Offices of other health practitioners	74.7	8,814	346	31,356	235.2	9,053
Individual & family services	21.1	9,411	478	11,919	65.3	1,096
Child day care services	39.2	9,747	691	14,125	516.6	6,263
Agents, managers for artists & public figures	1.8	502	10	1,549	25.0	867
Independent artists, writers & performers	8.9	2,272	38	4,461	465.1	9,631
Rooming & boarding houses	1.5	182	13	717	9.7	281
Full-service restaurants	91.9	20,908	1,863	47,763	29.0	3,308
Limited-service eating places	94.4	17,164	1,743	53,707	36.8	3,086
Special food services	11.2	8,314	563	17,466	68.5	2,014
Drinking places (alcoholic beverages)	44.3	3,753	360	12,468	21.3	1,466
Automotive repair & maintenance	94.8	12,533	583	46,945	251.2	11,570
Personal & household goods R&M	22.4	2,337	98	8,685	247.9	6,147
Personal care services	59.8	4,109	303	9,770	552.4	11,776
Drycleaning & laundry services	30.8	6,613	373	17,731	33.0	1,650
Other personal services	16.0	3,688	173	12,288	835.8	17,609
<b>Total Economy</b>	<b>5,443.40</b>	<b>3,773,003</b>	<b>113,658</b>	<b>16,871,471</b>	<b>15,536.07</b>	<b>711,264</b>

Firms and employment in thousands. Payroll and Revenue in millions.

**Table 4: Industry Revenue Shares and Business-Type Shares Within Industries, 2000**

NAICS Code	Industry Description	Industry Revenues, Percent of Aggregate Business Revenue	Industry Revenues, Percent Accounted for by Employers	Percent of Business Entities in Industry that Are Nonemployers
1152	Animal production support activities	0.01	31	96
2352	Painting & wall covering contractors	0.14	70	85
2355	Carpentry & floor contractors	0.30	68	88
2356	Roofing, siding, & sheet metal contractors	0.18	84	76
2357	Concrete contractors	0.12	88	63
3231	Printing & related support activities	0.25	97	54
3366	Ship & boat building	0.10	100	26
4471	Gasoline stations	1.08	99	15
4512	Book, periodical & music stores	0.08	93	75
4531	Florists	0.04	88	52
4853	Taxi & limousine service	0.04	50	96
4921	Couriers	0.11	99	32
5112	Software publishers	0.28	100	20
5242	Agencies & other insurance related activities	0.60	86	75
5312	Offices of real estate agents & brokers	0.40	67	91
5313	Activities related to real estate	0.32	68	90
5322	Consumer goods rental	0.07	94	54
5411	Legal services	0.92	93	58
5412	Accounting, tax prep, bookkeeping, payroll service	0.27	87	79
5415	Computer systems design & related services	0.86	94	77
5416	Management, scientific & tech consulting services	0.48	80	82
5615	Travel arrangement & reservation services	0.27	96	60
5617	Services to buildings & dwellings	0.40	84	82
6211	Offices of physicians	0.65	90	55
6212	Offices of dentists	0.12	91	46
6213	Offices of other health practitioners	0.23	78	76
6241	Individual & family services	0.07	92	76
6244	Child day care services	0.12	69	93
7114	Agents, managers for artists & public figures	0.01	64	93
7115	Independent artists, writers & performers	0.08	32	98
7213	Rooming & boarding houses	0.01	72	86
7221	Full-service restaurants	0.29	94	24
7222	Limited-service eating places	0.32	95	28
7223	Special food services	0.11	90	86
7224	Drinking places (alcoholic beverages)	0.08	89	32
8111	Automotive repair & maintenance	0.33	80	73
8114	Personal & household goods R&M	0.08	59	92
8121	Personal care services	0.12	45	90
8123	Drycleaning & laundry services	0.11	91	52
8129	Other personal services	0.17	41	98

**Table 5: Employer Links to Nonemployers by Size and Age of Employer, Employers in Selected Industries in 2000**

(1) Firm Size in 2000, Number of Employees	(2) Number of firms in size class	(3) Percent of payroll in size class	(4) Number of firms with backward links to nonemployers	(5) Percent of firms in size class with backward links	(6) Percent of all backward links to nonemployers	(7) Percent of nonemployer pre-link revenues
a)1-4	1416292	9.81	200252	14.14	75.36	46.73
b)5-9	432027	9.02	34590	8.01	13.02	13.09
c)10-19	242636	10.58	16656	6.86	6.27	12.38
d)20-49	139368	13.43	9229	6.62	3.47	8.62
e) 50-99	36886	8.56	2526	6.85	0.95	6.18
f)100-249	16426	8.83	1362	8.29	0.51	4.48
g)250-499	4332	7.16	489	11.29	0.18	2.32
h) 500+	3103	32.62	634	20.43	0.24	6.20
Total	2291070	100.00	265738	11.60	100.00	100.00

(1) Firm Age in 2000, Years Since First Appearance in Employer Universe	(2) Number of firms in age group	(3) Percent of payroll in age group	(4) Number of firms with backward links to nonemployers	(5) Percent of firms in age group with backward links	(6) Percent of all backward links to nonemployers	(7) Percent of nonemployer pre-link revenues
0	213292	2.28	35082	16.45	13.20	11.01
1	190690	3.27	31314	16.42	11.78	9.08
2	170091	3.40	27197	15.99	10.23	7.95
3	153400	3.44	25244	16.46	9.50	6.46
4	134315	3.40	20675	15.39	7.78	5.49
5	117723	3.10	20615	17.51	7.76	4.48
6 – 7	199523	5.61	24823	12.44	9.34	10.81
8+	1112036	75.50	80788	7.26	30.40	44.74
Total	2291070	100.00	265738	11.60	100.00	100.00

Note: Column (7) reports the pre-link percentage distribution of nonemployer revenues for nonemployers that link to firms in the employer universe. To calculate this distribution, we first express nonemployer revenues in 2000 dollars using the GDP deflator for all goods and services. Then, for each nonemployer that links to the employer universe, we take the value of its deflated revenue in the year prior to its first link to a firm in the employer universe. We sum these values across all nonemployers that link to the 2000 LBD for our selected industries. The percentages reported in the table are based on this total value of “year prior to link” nonemployer revenue.

**Table 6.A: Migrants from the Nonemployer Population as a Percentage of Young Employers (0 to 3 Years of Age) in 1997, Selected Industries**

		Percentage of Young Employers Accounted for by Migrants from the Nonemployer Population			
<i>First Appearance in the Nonemployer Population:</i>		1994 or earlier	1995	1996	All
SIC	Industry Description				
76	Farm labor and management services	24.5	13.1	15.8	53.5
78	Landscape and horticultural services	25.9	8.6	6.4	40.9
172	Painting and paper hanging	24.0	8.1	6.9	39.0
175	Carpentry and floor work	23.1	7.9	6.4	37.4
176	Roofing, siding, and sheet metal work	19.1	7.9	7.3	34.2
275	Commercial printing	13.4	5.6	4.8	23.8
367	Electronic components and accessories	9.8	4.3	6.4	20.4
412	Taxicabs	15.7	7.4	7.7	30.8
421	Trucking and courier services, except air	20.4	7.5	6.8	34.6
472	Passenger transportation arrangement	14.6	6.6	6.9	28.0
554	Gasoline service stations	8.4	3.6	3.5	15.6
581	Eating and drinking places	8.5	4.5	4.3	17.3
621	Security brokers and dealers	17.8	6.4	6.6	30.8
641	Insurance agents, brokers, and service	25.1	7.3	5.9	38.3
653	Real estate agents and managers	22.3	7.2	6.8	36.4
721	Laundry, cleaning, and garment services	17.9	6.5	5.3	29.7
723	Beauty shops	19.2	8.4	6.2	33.8
729	Miscellaneous personal services	22.2	8.4	7.2	37.8
737	Computer and data processing services	9.9	5.8	6.8	22.5
738	Miscellaneous business services	16.5	6.9	6.7	30.1
753	Automotive repair shops	19.0	7.4	5.5	31.9
784	Video tape rental	16.0	7.1	6.3	29.5
792	Producers, orchestras, entertainers	15.5	6.2	6.6	28.2
799	Misc. amusement, recreation services	12.3	5.4	5.4	23.1
801	Offices and clinics of medical doctors	11.4	4.3	4.5	20.2
802	Offices and clinics of dentists	17.2	6.6	4.9	28.7
803	Offices of osteopathic physicians	14.2	5.0	4.5	23.7
804	Offices of other health practitioners	18.4	6.8	5.9	31.1
811	Legal services	18.1	6.4	4.9	29.4
832	Individual and family services	10.6	4.0	3.7	18.3
835	Child day care services	21.9	8.0	6.0	35.8
872	Accounting, auditing, and bookkeeping	21.8	6.4	5.1	33.2
874	Management and public relations	12.3	5.6	6.6	24.5
	All Selected Industries	16.1	6.3	5.7	28.1



**Table 6.B: Migrants from the Nonemployer Population as a Percentage of Young Employer Revenues in 1997, Selected Industries**

		Percentage of Young Employer Revenues Accounted for by Migrants from The Nonemployer Population			
<i>First Appearance in the Nonemployer Population:</i>		1994 or earlier	1995	1996	All
SIC	Industry Description				
76	Farm labor and management services	20.3	18.9	19.2	58.4
78	Landscape and horticultural services	19.3	5.8	4.6	29.7
172	Painting and paper hanging	17.7	6.3	6.1	30.0
175	Carpentry and floor work	16.6	7.0	5.7	29.2
176	Roofing, siding, and sheet metal work	13.1	5.3	6.4	24.7
275	Commercial printing	7.3	3.7	4.1	15.0
367	Electronic components and accessories	4.0	1.0	3.3	8.3
412	Taxicabs	18.2	7.3	9.0	34.5
421	Trucking and courier services, except air	16.6	7.9	7.1	31.6
472	Passenger transportation arrangement	12.0	5.1	6.8	23.9
554	Gasoline service stations	6.5	2.8	2.6	11.8
581	Eating and drinking places	5.6	3.0	2.8	11.4
621	Security brokers and dealers	9.7	4.2	5.6	19.5
641	Insurance agents, brokers, and service	16.0	5.0	6.0	27.1
653	Real estate agents and managers	18.6	10.8	8.9	38.3
721	Laundry, cleaning, and garment services	11.9	5.0	3.4	20.3
723	Beauty shops	15.4	6.4	4.8	26.5
729	Miscellaneous personal services	15.3	15.5	6.1	36.9
737	Computer and data processing services	8.2	4.1	5.7	17.9
738	Miscellaneous business services	12.1	4.9	7.7	24.7
753	Automotive repair shops	12.0	5.2	3.4	20.6
784	Video tape rental	10.9	4.3	4.1	19.3
792	Producers, orchestras, entertainers	9.8	4.8	6.8	21.4
799	Misc. amusement, recreation services	8.4	5.7	4.4	18.4
801	Offices and clinics of medical doctors	7.9	3.4	4.6	15.9
802	Offices and clinics of dentists	13.8	5.1	3.8	22.8
803	Offices of osteopathic physicians	10.6	3.0	2.1	15.7
804	Offices of other health practitioners	12.9	5.3	5.1	23.2
811	Legal services	11.5	4.2	3.7	19.3
832	Individual and family services	5.9	2.6	4.1	12.7
835	Child day care services	12.8	4.5	4.1	21.4
872	Accounting, auditing, and bookkeeping	12.5	3.7	7.0	23.2
874	Management and public relations	7.7	4.7	7.3	19.8
	All Selected Industries	10.4	4.8	5.2	20.4

Note: “Young Employers” in 1997 are businesses that first hire one or more paid employees in 1995, 1996 or 1997. In this sense, they are 0 to 3 years of age as of 1997.

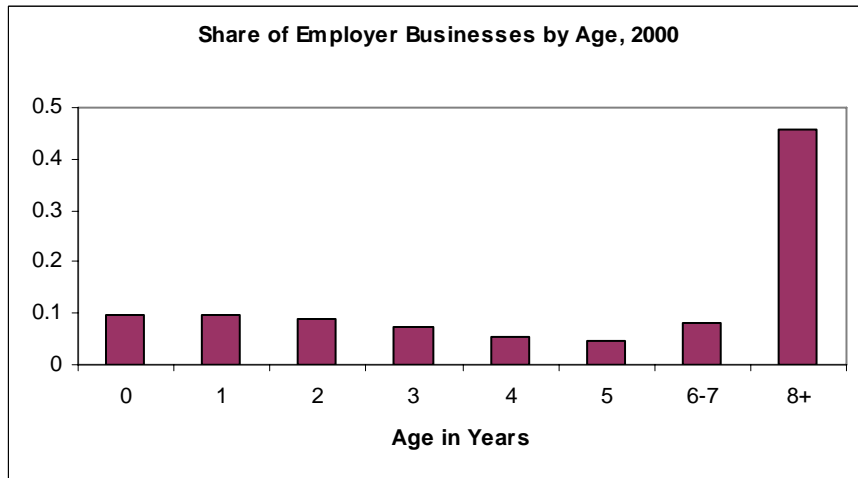
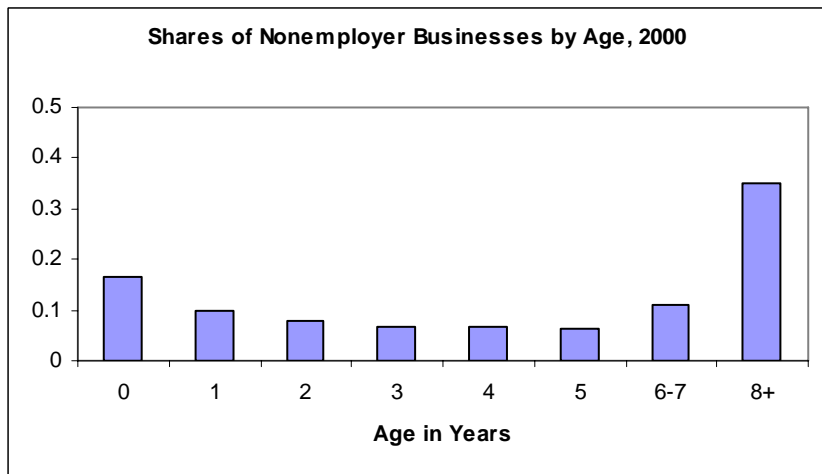
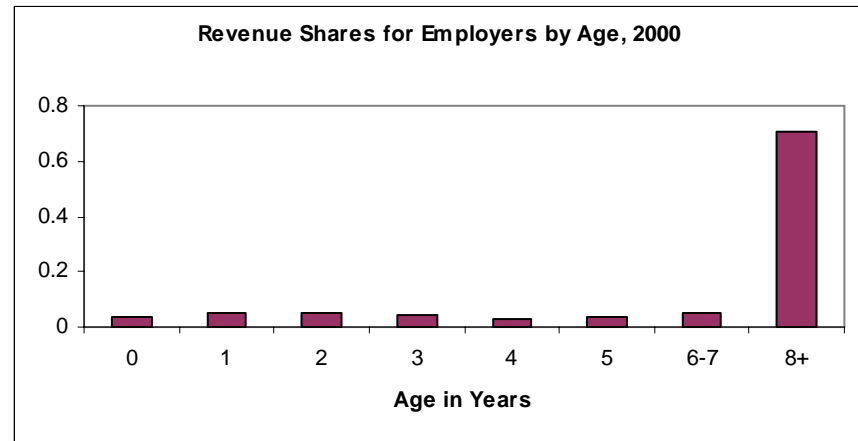
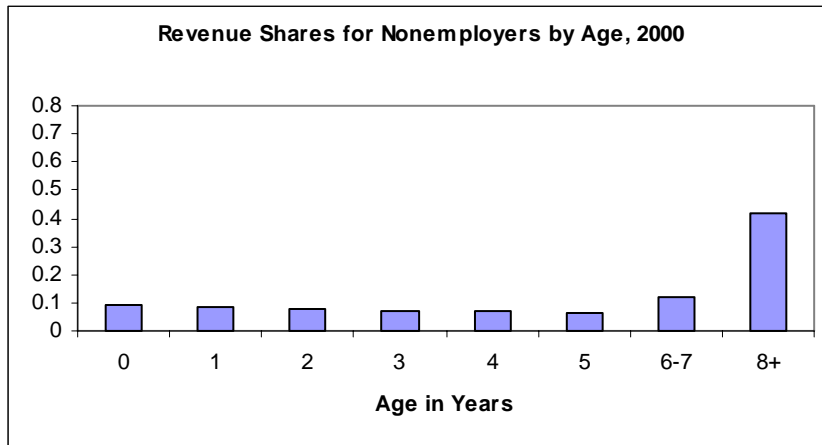
**Table 7: Summary Statistics for Nonemployer Revenue Growth Rates by Transition Status**

	Time Interval	Migrants	Continuers and Exits	Continuers Only
<b>Mean</b>	95-96	31%	23%	20%
<b>Median</b>	95-96	15%	5%	5%
<i>10<sup>th</sup> percentile</i>	95-96	-26%	-51%	-44%
<i>90<sup>th</sup> percentile</i>	95-96	141%	193%	139%
<b>90-10 Differential</b>	95-96	167%	244%	183%
<b>Mean</b>	96-97	101%	-14%	6%
<b>Median</b>	96-97	102%	0%	3%
<i>10<sup>th</sup> percentile</i>	96-97	-5%	-197%	-59%
<i>90<sup>th</sup> percentile</i>	96-97	200%	70%	76%
<b>90-10 Differential</b>	96-97	205%	267%	135%

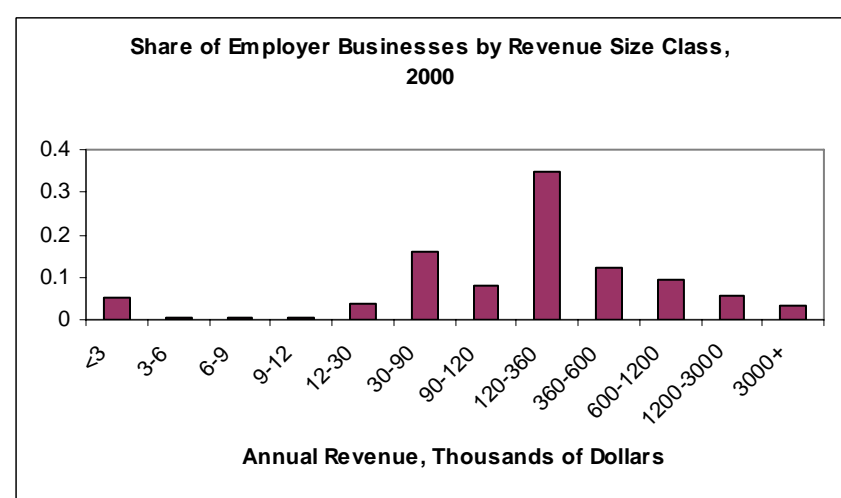
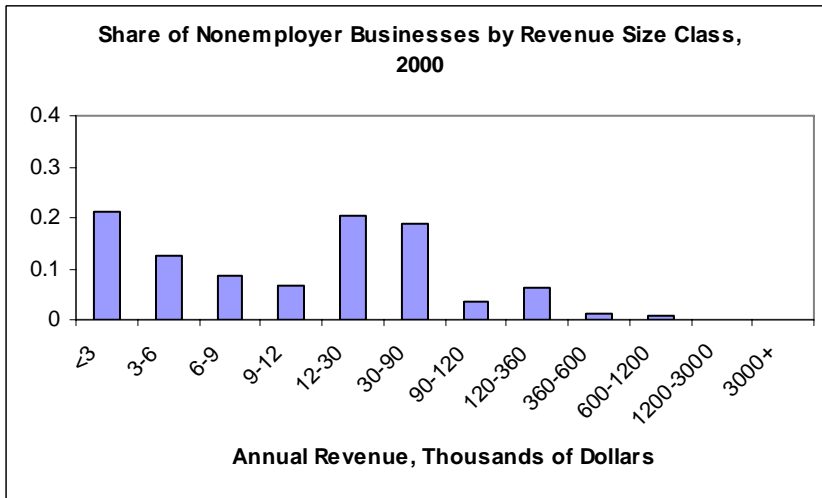
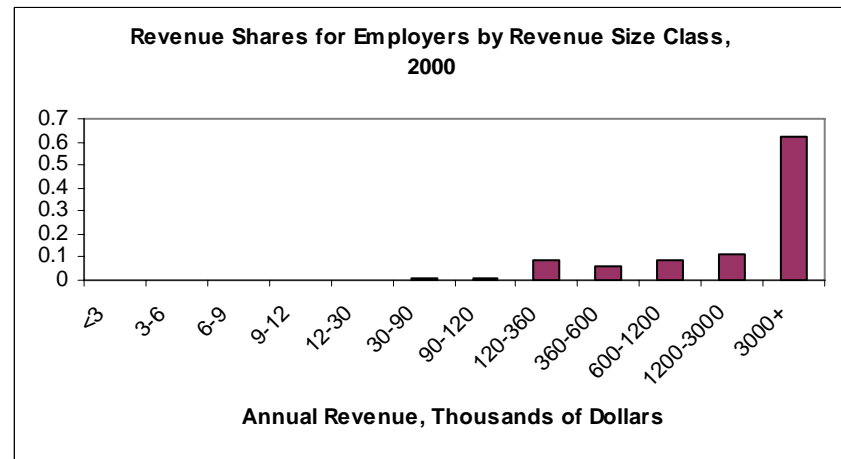
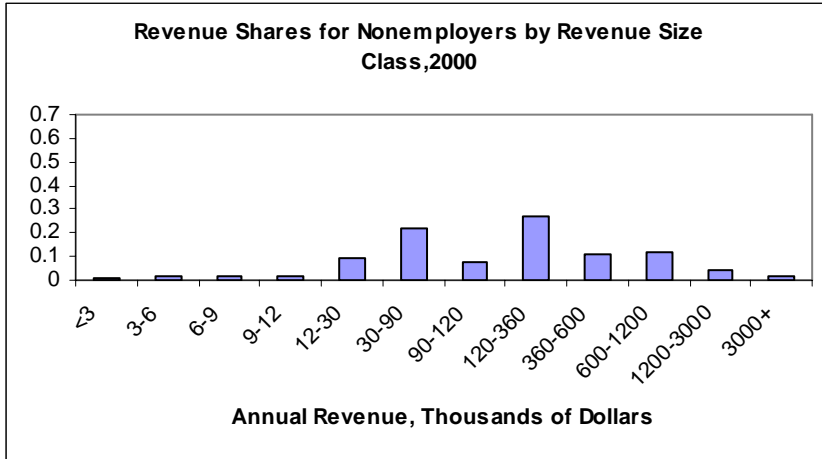
Notes:

1. The analysis population contains all nonemployers in our selected industries with positive revenue in 1996 that are classified as either Migrants, Continuers or Exits based upon their 1996-97 transitions.
2. Table entries report summary statistics for the distribution of annual revenue growth rates from 1995 to 1996 and from 1996 to 1997. The revenue growth rate is measured as the change in annual revenue from  $t - 1$  to  $t$  divided by the simple average of revenue in  $t - 1$  and  $t$ . All statistics are computed on a revenue-weighted basis.

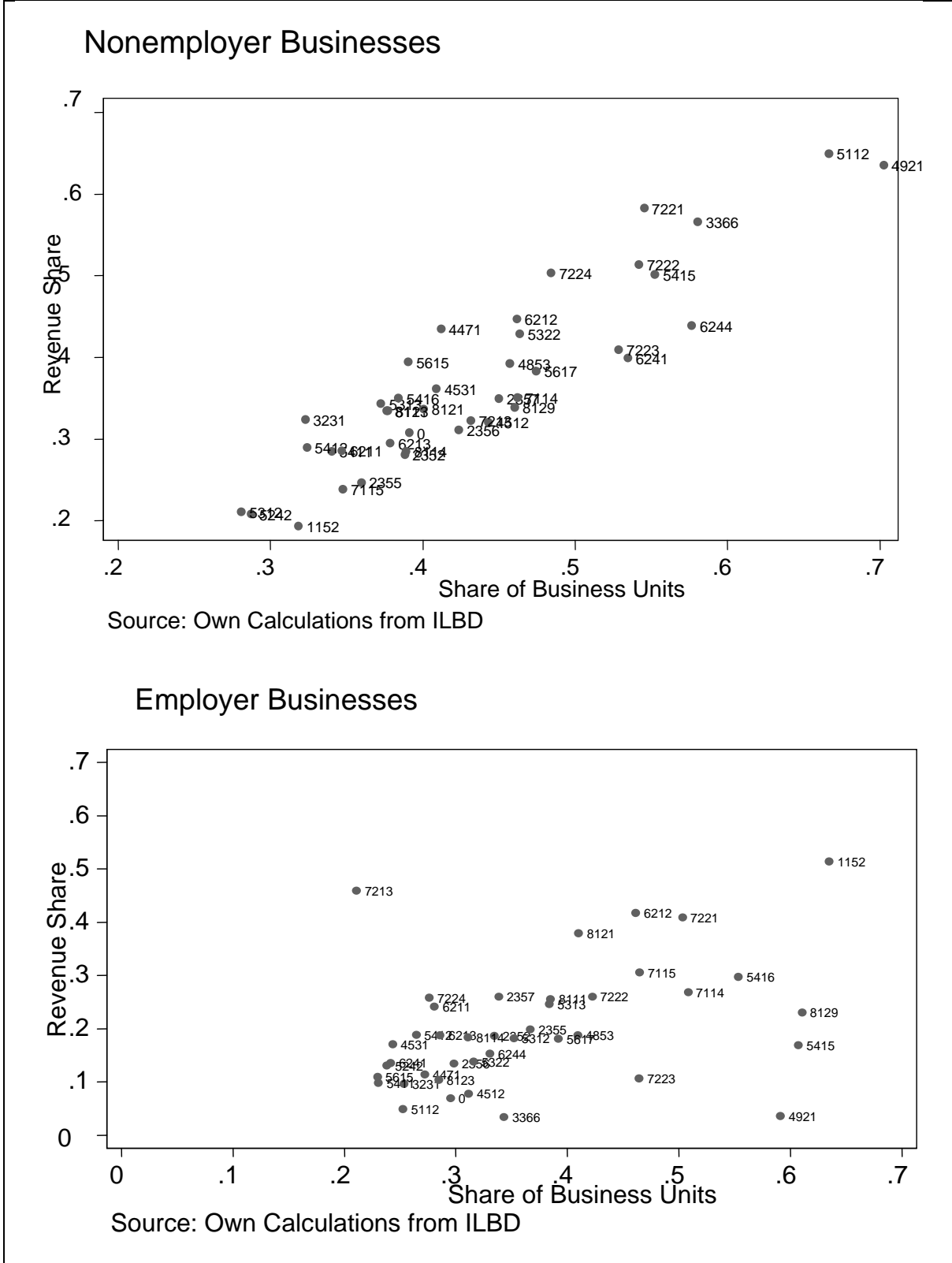
**Figure 1: Age Distribution of Business Numbers and Revenues within Each Universe**



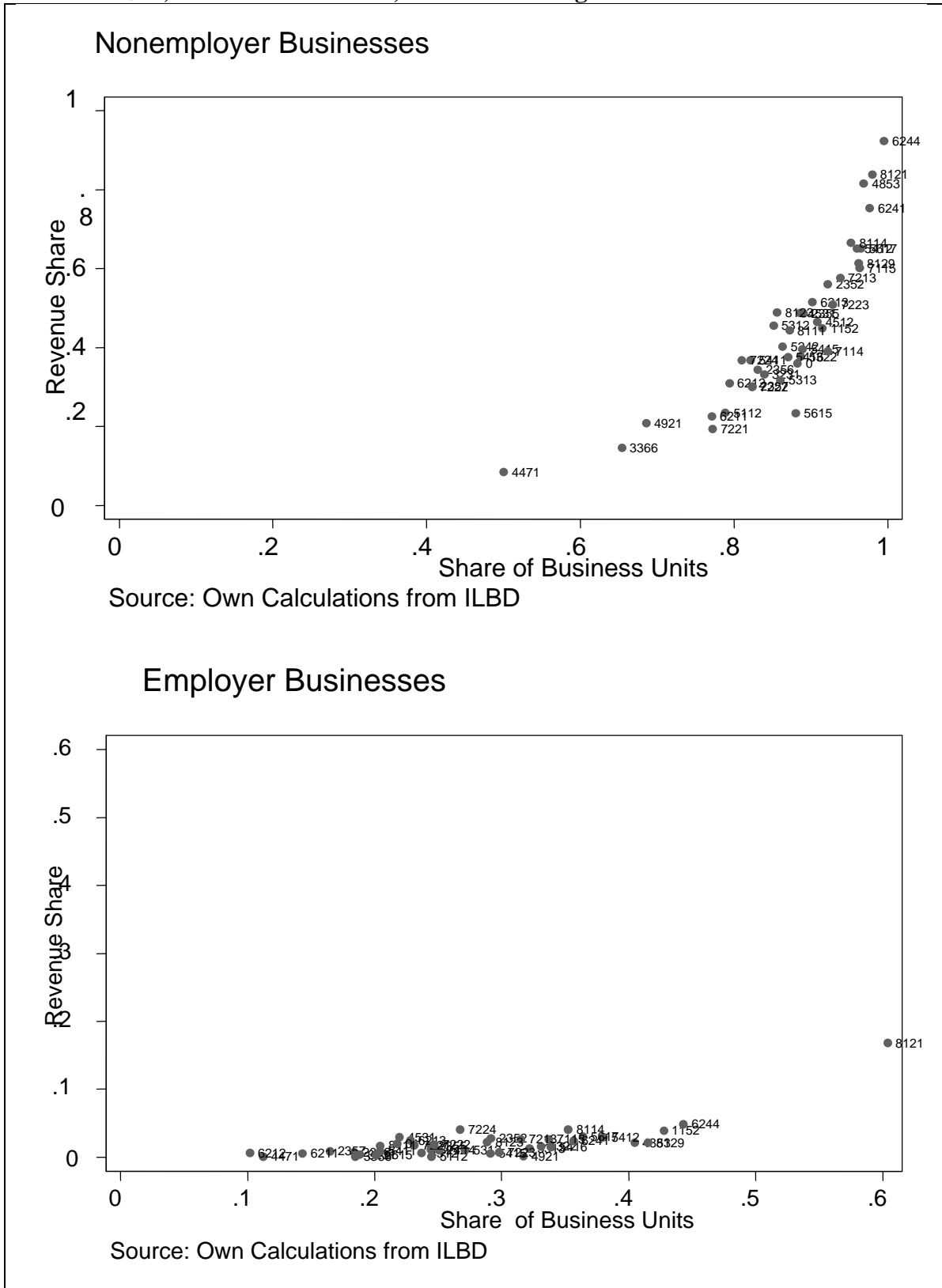
**Figure 2: Size Distribution of Business Numbers and Revenues with Each Universe**



**Figure 3: Share of Revenues and Business Numbers Accounted for by Young Firms, 0 to 3 Years Old, in Selected 4-Digit NAICS Industries**



**Figure 4: Share of Revenues and Business Numbers Accounted for by Small Firms, Less than \$90,000 Annual Revenue, in Selected 4-Digit NAICS Industries**



**Figure 5: Age Difference between Matched Employer and Nonemployer Records**

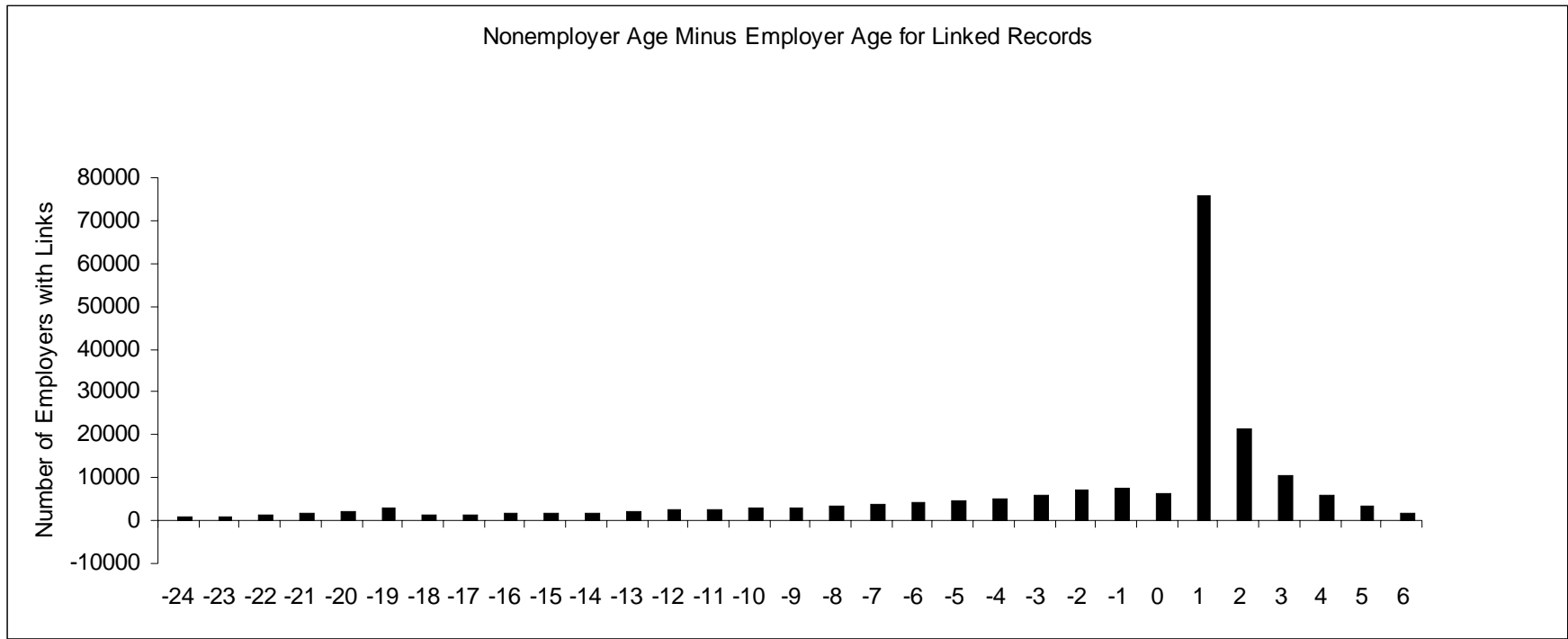
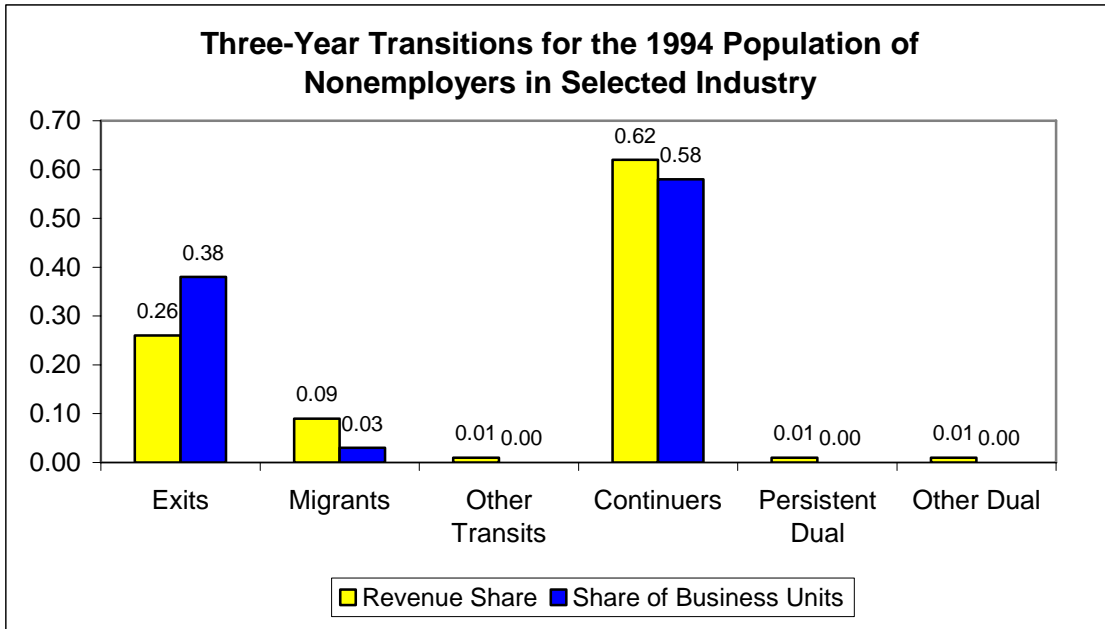
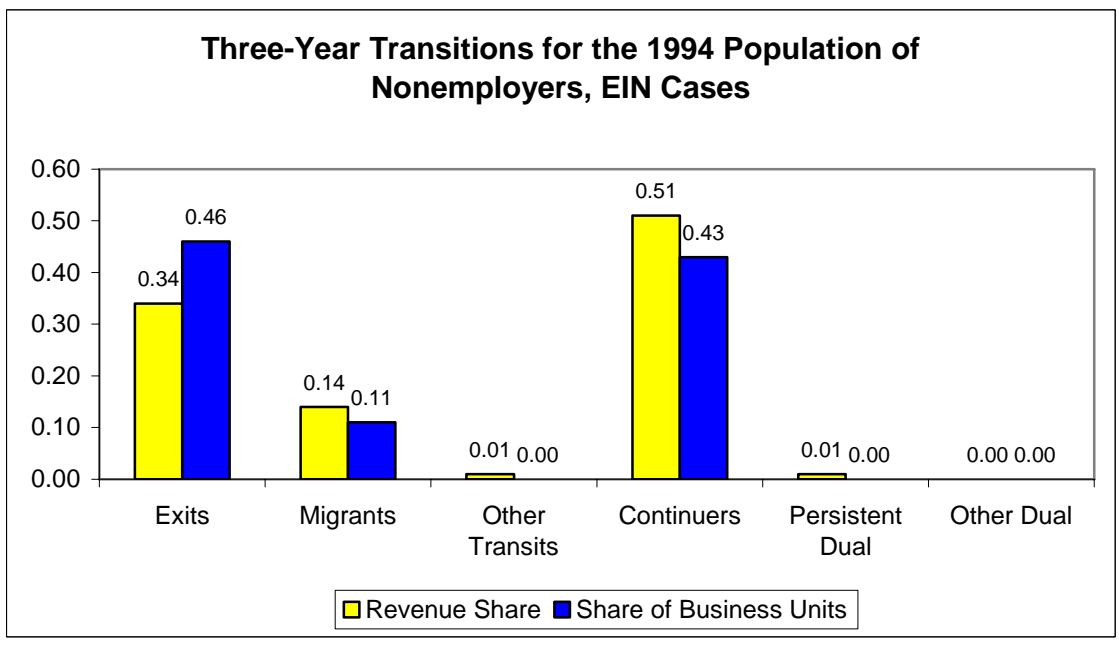
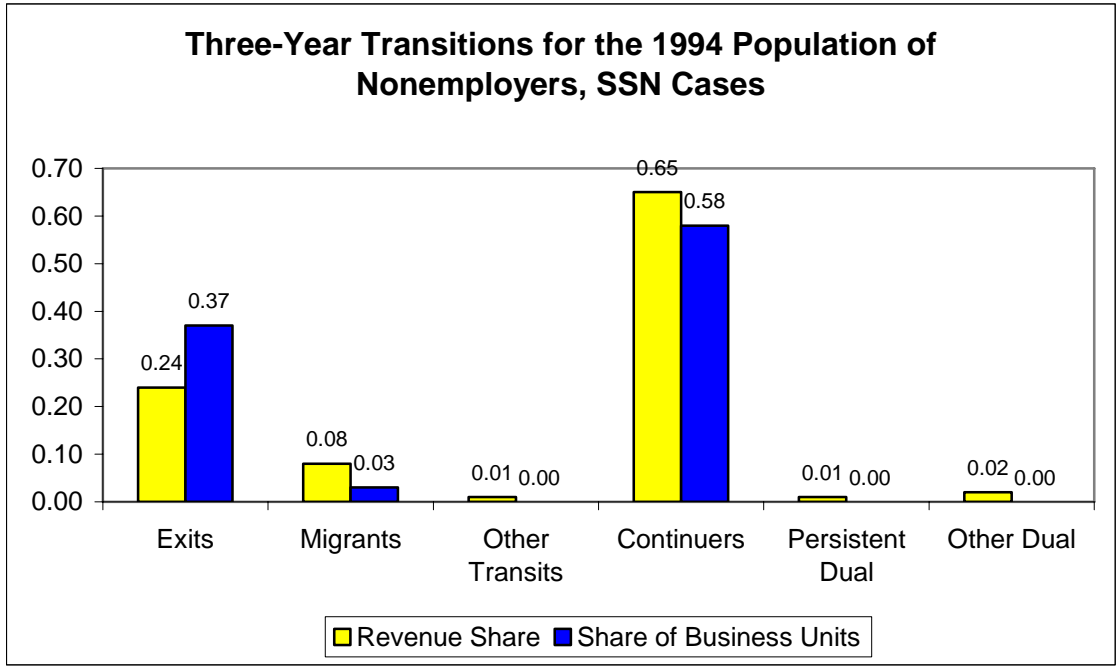


Figure 6

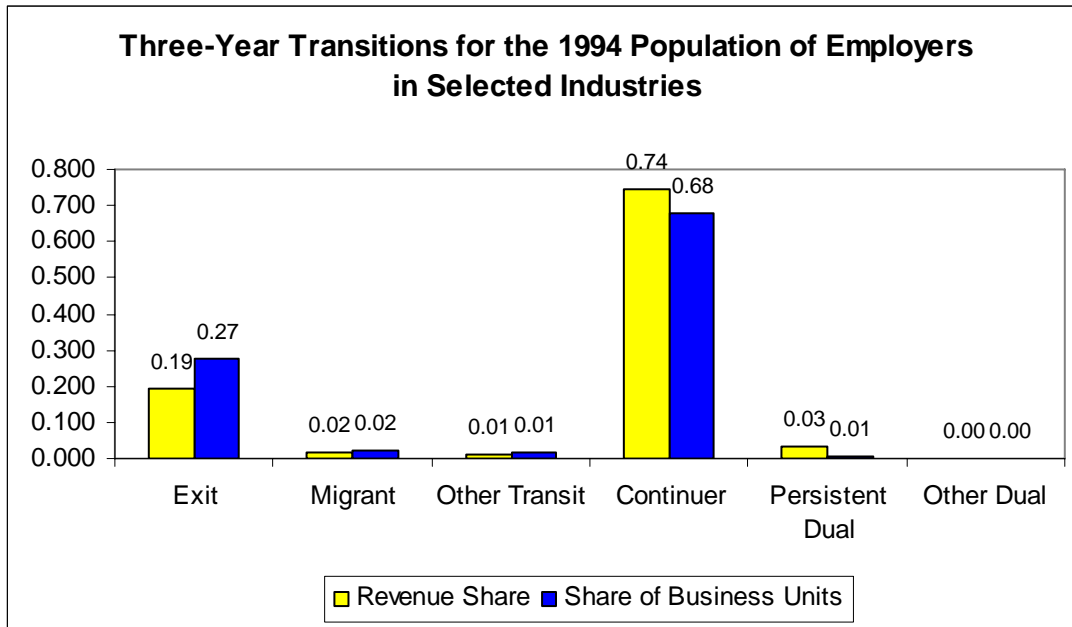




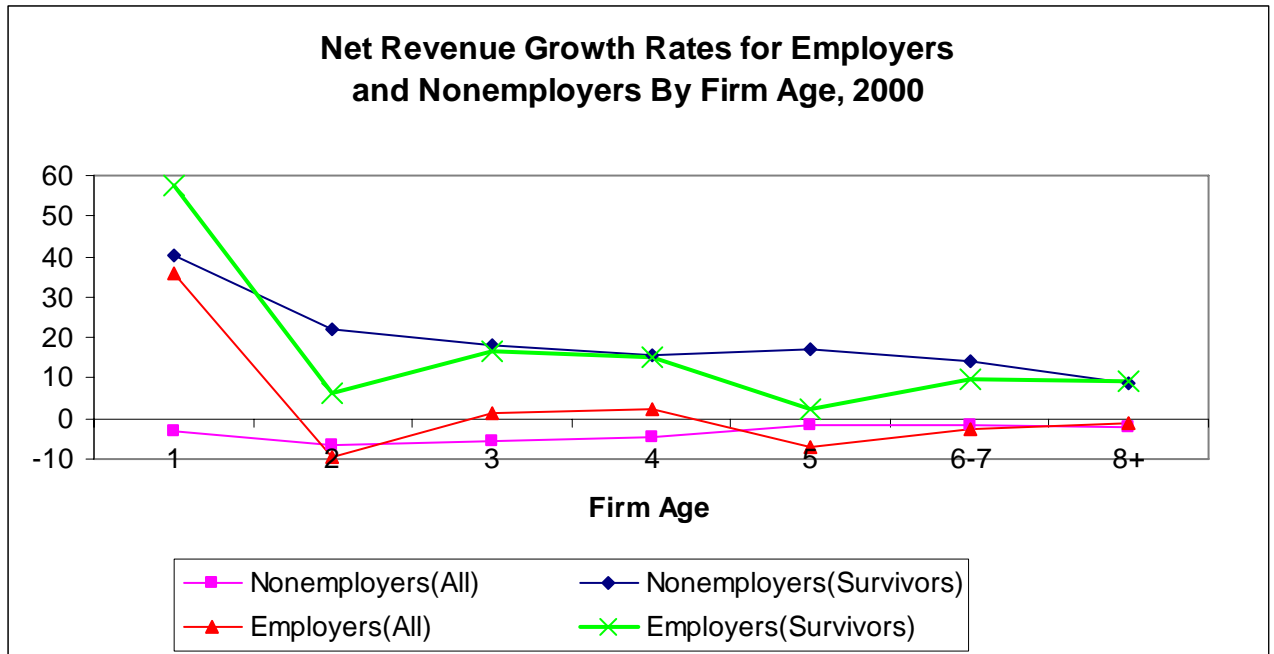
**Figure 7**



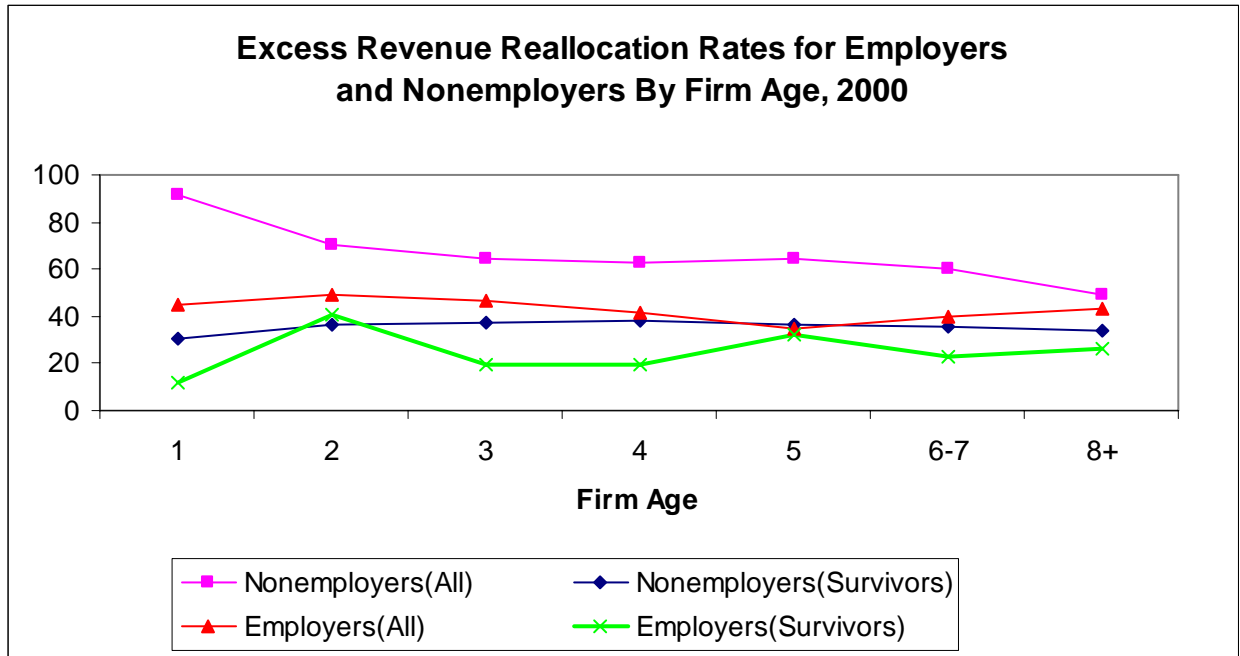
**Figure 8**



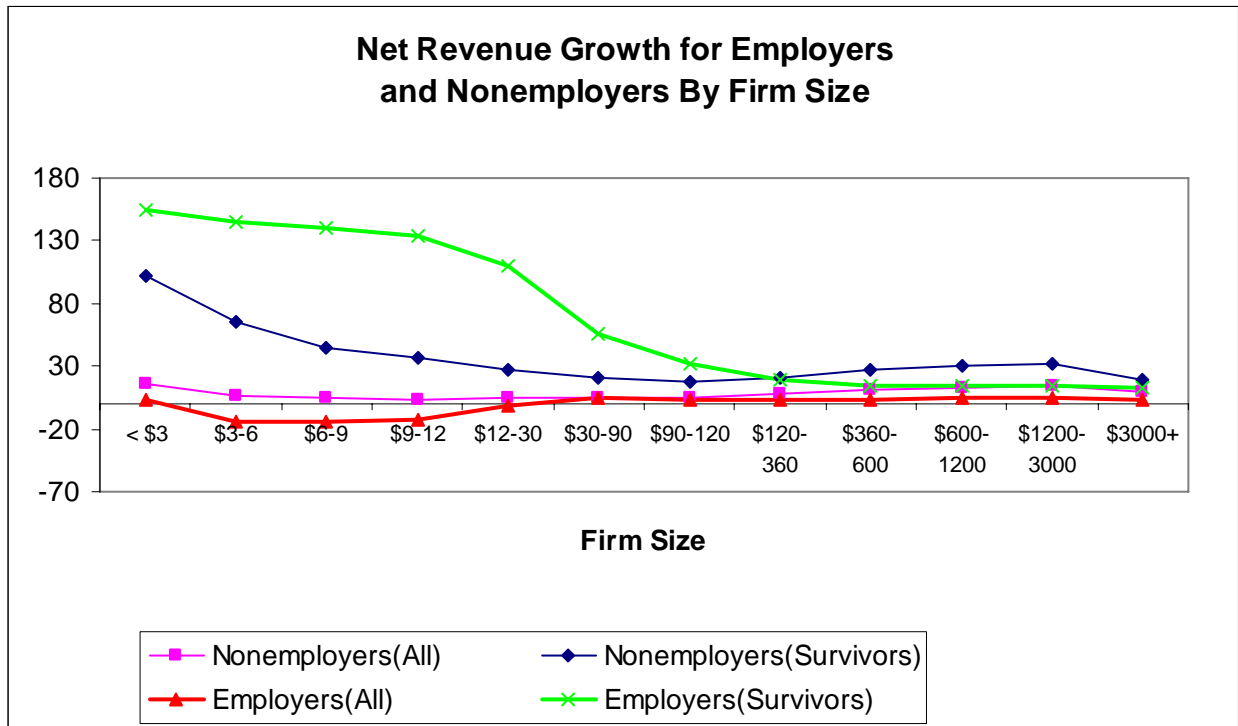
**Figure 9: Annual Revenue Growth by Business Type and Age**



**Figure 10: Excess Revenue Reallocation by Business Type and Age**



**Figure 11: Mean Revenue Growth by Business Type and Size**



**Figure 12: Excess Revenue Reallocation by Business Type and Size**

