Evolution of Firm Dynamics: With Minimal Assumptions*

Kim P. Huynh† David T. Jacho-Chávez‡ Robert J. Petrunia§ Marcel Voia¶

August 24, 2009

Abstract

This paper investigates the evolution of firm distributions for entrant manufacturing firms in Canada using nonparametric methods. These nonparametric methods allow a flexible method based on functional principal components or dynamic densities to characterize how these densities evolve over time. This method is applied to a novel administrative firm-level database from Statistics Canada to investigate the evolution of the 1985 and 1989 cohorts of new entrants. We find that firm leverage (debt-to-asset ratio) distributions have persistent deviations from the initial distributions and bootstrap test statistics suggest that the distributions are different across all time periods. Firm size and labour productivity have transitory deviations and some of the distributions are the same across all time periods. Univariate finite mixture and stochastic dominance tests are used to conduct pairwise comparisons as robustness measures. We find that these static pairwise comparisons confirm the dynamic evolution of these densities. This method illustrates the efficacy of functional principal components to analyze firm distributions.

Keywords: Firm Distributions, Functional Principal Components, Nonparametric Methods.

---

*The assistance and hospitality of Statistics Canada Business and Labour Market Analysis and the Economics Analysis department is gratefully acknowledged. We are especially grateful to Garnett Picot for his support and encouragement in this project. We thank Victor Aguirregabiria and the participants of 2009 Canadian Economics Association. We are indebted to Steve Romer for his help with computing issues. The contents of this paper have been subject to vetting and pass the Disclosure Rules & Regulations set forth by Statistics Canada. We thank Susan Brunet and Bob Gibson for vetting and sending our results expeditiously. All remaining errors are the responsibility of the authors.

†Department of Economics, Indiana University, 105 Wylie Hall, 100 S Woodlawn, Bloomington, IN 47405, USA. Phone: +1 (812) 855 2288. Fax: +1 (812) 855 3736. E-mail: kphuynh@indiana.edu.

‡Department of Economics, Indiana University, 105 Wylie Hall, 100 S Woodlawn, Bloomington, IN 47405, USA. Phone: +1 (812) 855 728. Fax: +1 (812) 855 3736. E-mail: djachoch@indiana.edu.

§Department of Economics, Lakehead University, 955 Oliver Road, Thunder Bay, ON P7B 5E1, CANADA. Phone: +1 (807) 343 8384. Fax: +1 (807) 346-7936. Email: rpetruni@lakeheadu.ca

¶Department of Economics, Carleton University, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, CANADA. Phone: +1 (613) 520-2600 x3546. Fax: +1 (613) 520-3906. Email: mvoia@connect.carleton.ca

1 Introduction

Understanding the evolution of firm distributions has been a major concern among economists, businesses, and policymakers. This paper investigates the evolution of firm debt-to-asset ratio (leverage), labour productivity, and size distributions of manufacturing firms in Canada. Previous literature has focused mostly on productivity and firm size distributions. The novelty of this analysis is the inclusion of financial information (debt-to-asset ratio) of small, young, and privately-owned firms. However, little is known about the evolution of financial variables due to the lack of detailed financial data on small, private firms. This paper uses the T2LEAP, a unique administrative firm-level database combining information from the Corporate Tax Statistical Universe File (T2SUF) with Longitudinal Employment Analysis Program (LEAP) database in Canada. Information in the database allows us to analyze the evolution of debt-to-asset ratio distributions, in addition to size and productivity distributions, for two entry cohorts (1985 and 1989) in the manufacturing sector. The analysis focuses on nonparametric methods to examine variable distributions from both a static and dynamic perspective.

Beginning with Gibrat (1931), firm size distributions (FSD) have been a concern for studies on industry dynamics. Typically, FSD for an industry are found to be skewed and lognormal in shape. Work on firm dynamics suggests financial frictions are relevant in the evolution of firm size distributions. Cabral and Mata (2003) document that the FSDs move towards a lognormal distribution with age. One possible explanation for the changes in FSD associated with age is a learning and selection process as suggested in Jovanovic (1982). Cabral and Mata (2003) provide evidence against selection and suggest a simple model where the relaxing of financial constraints with firm age generate the dynamics observed for FSD. Angelini and Generale (2008) provide further evidence of the impact of financial constraints on the evolution of FSD. Recent work using the T2LEAP by Huynh and Petrunia (2008) documents empirically that a firm's growth depends nonlinearly on its leverage (debt-to-asset ratio) even when controlling for firm size and age. Huynh, Petrunia, and Voia (forthcoming) find a similar nonlinear conditional relationship between a firm's survival prospects and its initial leverage; survival rates increase with leverage for firms in the low to mid range of the leverage distribution and fall for firms at the upper end.

We perform the following procedures to capture both the static and dynamic statistical properties of variable distributions:

1. Functional principal component analysis (FPCA) is used to perform dynamic analysis of variable distributions for each entry cohort. Debt-to-asset ratio sees persistent changes to a cohort's distribution over time. Movements of the size and productivity distributions are mostly transitory.

2. The equality of densities of each variable's distribution across years is tested via a bootstrap procedure. The bootstrap procedure analyzes distributions over time for entry cohorts between 1985 and 1993. Across time debt-to-asset ratio distributions are found to be not jointly equal statistically for all cohorts. Joint equality of an entry cohort's across time productivity distributions is rejected for every cohort except the 1990 entry cohort. We do not reject equality in across time size distributions for the 1985, 1991, 1992, and 1993 entry cohorts, while joint equality is rejected for all other entry cohorts.

3. Pairwise comparisons of distributions are made using stochastic dominance tests and finite mixture analysis. Finite mixtures are more stable for size and productivity distribution, while continual movements occur for the mixtures of the debt-to-asset ratio distribution. Stochastic dominance tests indicate the debt-to-asset ratio and size distributions change with an entry cohort's age, while no movement occurs for a cohort's productivity distribution.

The rest of the paper is organized as follows: Section 2 provides a description of the data used in the paper. Section 3 discusses the methodology and the results of this dynamic distributional analysis. Pairwise comparison of the cross-sectional aspects of variable distributions are examined in Section 4. Finally, section 5 concludes.

2 T2LEAP: Firm-Level Data

2.1 Data Description

The firm-level data used in this study comes from the T2LEAP database maintained by Statistics Canada. This database was created through the merging of two administrative databases; employment information from the Longitudinal Employment Analysis Program (LEAP) is linked to financial records from the Corporate Tax Statistical Universe File (T2SUF). A firm is incorporated if it files a corporate (T2) tax return. T2LEAP uses a business registry number (BSNUM) to track all incorporated firms operating in a given year. The data effectively covers the universe of incorporated Canadian firms hiring workers. The T2LEAP database contains firm details from 1984 until 1997. A firm's entry into the database allows us to identify birth year and place firms into entry cohorts. Birth year is measured as the first year in which a firm both hires employees and files a corporate tax return. We do not observe a birth year for firms existing in the database in 1984. A firm is considered to exit when it fails to hire employees and/or file a tax return. The removal of the firm occurs in all years subsequent to exit. Each firm receives an unique identify, BSNUM, which ensures that exit and re-entry does not occur.

The paper uses the following firm level information available in the T2LEAP database: Annual measures of a firm's employment, sales, assets, and debt. Each firm is classified by a three-digit Standard Industry Classification (SIC) code. The paper only considers manufacturing firms. The following generated variables are calculated using the firm information: Leverage or the debt-to-asset ratio ($\text{Debt}_{it}/\text{Asset}_{it}$), and labour productivity or ($\text{Sales}_{it}/\text{ALU}_{it}$), where the Average Labour Units (ALU) measure a firm's yearly employment. Analysis is performed on leverage, the logarithm of sales and labour productivity.

2.2 Univariate Densities

We perform the analysis on two entry cohorts: (i) 1985 and (ii) 1989. Focusing on new entrants allows us to capture the dynamic process related to the post-entry evolution of new firms. The 1985 entry cohort is the first identifiable birth cohort, and thus, is the entry cohort with the longest time-series. The 1989 entry cohort provides a comparison group, but also allows for us to separate age effects from time effects. The 1989 cohort is chosen since 1989 is the first year after the Canada-US free trade agreement (CUSFTA) and last year prior to the recession of early 1990s.

Figures 1 illustrates the densities for leverage, logarithm of sales, and labour productivity for the two entry (1985 and 1989) cohorts. In each figure there is a density for ages one, four and seven. In the first row, the leverage distribution shows a clear pattern of moving leftwards left with age. As firms age (survive) they reduce their debt relative the asset holdings. For the 1989 entry cohort, there is clear decrease in leverage with age as the leverage distribution shifts right with age. Alternatively, the decrease is not uniform as a crossing occurs between the age four and age seven leverage distributions for the 1985 cohort. The results are not surprising since at age seven the 1985 cohort is in the midst of a deep recession in Canada while age seven for the 1989 would have seen a recovery.

The second row contains the logarithm of sales, which is our measure of firm size. The distinguishing feature of firm size is that the densities shift to the right as the entrant cohort ages. This stylized fact is known as the 'trend to bigness', see Lucas (1978). The difference between the two cohorts is that the 1989 has more mass in terms of larger size firms while 1985 has smaller firms.

Finally, row three contains the densities for labour productivity. In contrast to the two other variables, the labour productivity distribution does not show as much movement much with age. The 1989 cohort has higher productivity and has less dispersion in labour productivity than the 1985 cohort. Overall, the labour productivity for the most part increases with age for both cohorts.

Based on these observations, we note that as firms age they reduce their leverage, get larger, and more productive. Between cohort comparison shows that members of the 1989 cohort reduce their leverage faster, are larger in size, and are more productive. As shown in Huynh, Petrunia, and Voia (forthcoming), the over time changes to sales and leverage distributions likely result from two factors: (i) selection due to exit of firms with age and (ii) changes from initial conditions for surviving firms. The next section describes a general method to describe the evolution of these densities in a formal statistical fashion.

3 Dynamic Distribution Analysis

We make use of functional principal component analysis (hereafter FPCA), as suggested by Kneip and Utikal (2001), to describe the underlying population densities of each cohort and age group jointly in labour productivity, leverage, and firm size distributions. To the best of our knowledge, the usage of this tool is novel in understanding firm distributions and industrial dynamics. It is a powerful device that is able to decompose the changes in distributions into dynamic factors. In the time-series and panel data literature Bai and Ng (2002) and Pesaran (2006) have proposed to use factor models. The difference lies in how the factors are estimated. We focus on nonparametric kernel methods since we want to place minimal structure on how the dynamic factors are computed. The next section presents these procedures.

3.1 Functional Principal Component Analysis

The approach of Kneip and Utikal (2001) analyzes jointly the underlying population densities $\{f_t\}_{t=1}^{T}$. In particular, to characterize differences and similarities of $\{f_t\}_{t=1}^{T}$, we assume their expansions into the first L principal components, $g_1, g_2, \ldots, g_L$, and represent each f_t in terms of the model

$$f_t = f_\mu + \sum_{j=1}^{L} \theta_{tj} g_j, \tag{1}$$

where $f_\mu = \sum_{t=1}^{T} f_t/T$ is the common mean and $L \leq T$ corresponds to the number of nonzero eigenvalues of the empirical covariance operator.[1] Model (1) implies that each f_t can be

[1] Consider the space of square-integrable functions ξ such that $\int \xi^2(x)\, w(x)\, dx < \infty$, where $w(x) > 0$ is some continuous, uniformly bounded weighting function $\forall x \in D \subset \mathbb{R}$ that lies in the support of f_t, $\forall t$; and define the scalar product $\langle \xi_1, \xi_2 \rangle = \int \xi_1(x)\, \xi_2(x)\, w(x)\, dx$ and $\|\xi\|^2 = \langle \xi, \xi \rangle$. Then the empirical covariance operator of $\{f_t\}_{t=\tau}^{T}$ is given by $\mathrm{V}\xi = (1/T) \sum_{t=\tau}^{T} \langle f_t - f_\mu, \xi \rangle (f_t - f_\mu)$, where τ is the start year and T is the last observed year. In this paper we use the uniform weight function, i.e. $w(x) = 1$, $\forall x \in D$.

obtained by adding to f_μ a transformation of compromising common components g_1, g_2, $\ldots$, g_L, with varying strengths encapsulated in the coefficients θ_{tj}. Since f_t represent densities obtained for each time period $t = 1, \ldots, T$, then the time evolution of their respective coefficients θ_{t1}, θ_{t2}, $\ldots$, θ_{tL} provides information about the evolution of the main differences and similarities between the underlying distributions.

The unknown g_1, g_2, $\ldots$, g_L and θ_{t1}, θ_{t2}, $\ldots$, θ_{tL}, can be obtained from the $T \times T$ matrix $\mathbf{M}$, whose elements are defined by $\mathbf{M}_{ts} = \langle f_t - f_\mu, f_s - f_\mu \rangle$, $\forall t, s = 1, \ldots, T$. In particular, the unknown components g_r, and parameters θ_{tr} relate to the T eigenvectors, $\mathbf{p}_r = (p_{1;r}, \ldots, p_{T;r})^\top$, $r = 1, \ldots, T$, with corresponding eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots, \geq \lambda_T$ of $\mathbf{M}$ through:

$$\theta_{tr} = \lambda_r^{1/2} p_{t;r},\ g_r = \lambda_r^{-1/2} \sum_{t=1}^{T} p_{t;r}, \text{ and } f_t = \frac{\sum_{t=1}^{T} \theta_{tr} f_t}{\sum_{t=1}^{T} \theta_t^2}. \tag{2}$$

Kneip and Utikal (2001) provide consistent estimators of θ_{tr},[2] and g_r. The next section describes the results from these FPCA estimates.

3.2 Test of Densities via Bootstrap

The data contains the universe of all firms for the period of 1984-1997. In essence there are 13 cohorts, 12 of which have a known entry date while those in the 1984 have censored entry date. The first step is to construct a test of whether the densities of these variables come from the same distribution, i.e. $H_0 : \ L = 0$ as suggested in (Kneip and Utikal 2001). They proposed using $\widehat{\rho} = \sum_{r=1}^{T} \widehat{\lambda}_r$ as a plausible test statistic, and the following bootstrap procedure for approximating its distribution: In each of 399 replications T subsamples of sizes $n_1^*, \ldots, n_T^*$ are drawn with replacement from the pooled empirical cdf of the entire

[2] However, no asymptotic distribution was derived.

sample, i.e. $\sum_{t=1}^{T} n_t$, and calculated $\widehat{\rho}^* = \sum_{r=1}^{T} \widehat{\lambda}_r^*$. The null distribution of the test statistic is then approximated by the empirical distribution of $\widehat{\rho}^*$. The results are shown in Table 1. The bootstrap test statistic for the leverage distribution rejects the null hypothesis of equal distributions across all cohorts at the 10% significance level. The result for size and labour productivity is different as there are some cohorts that does not reject the null hypothesis of equal distributions. Namely, the 1975, 1991, 1992, and 1993 cohorts for size while 1990 cohort for labour productivity.

Table 1: Bootstrap Test Statistics and P-values

Cohort	Leverage	p-value	Size	p-value	Labour Productivity	p-value
84	476.422	0.000	13.440	0.028	147.856	0.000
85	205.937	0.000	11.610	0.318	97.762	0.000
86	193.368	0.000	11.391	0.015	86.358	0.020
87	167.574	0.000	9.376	0.003	43.955	0.005
88	146.996	0.000	7.973	0.000	47.092	0.045
89	123.089	0.000	5.777	0.005	43.152	0.045
90	101.145	0.000	4.528	0.010	39.214	0.278
91	86.000	0.020	3.139	0.667	28.966	0.098
92	71.180	0.093	2.621	0.233	14.420	0.050
93	55.195	0.090	2.070	0.419	10.975	0.015

Note: P-values calculated based on 399 bootstrap replications in each cohort.

3.3 Dynamic Scree Plots

Figure 2 displays $\widehat{\lambda}_r / \sum_{r=1}^{T} \widehat{\lambda}_r$, where $\widehat{\lambda}_r$ represent the estimated eigenvalues of the FPCA decompositions of the 1985 and 1989 cohorts across the three variables. The estimated eigenvalues are plotted across time and borrowing from the principal components literature we call these *Dynamic Scree Plots* to emphasize the time-series dimension of these objects. For leverage the dynamic scree plots show that the eigenvalues do not decline until the final time period at which it falls to zero by definition. For firm size there is a marked decline in the eigenvalues. Labour productivity shows a decline and then plateau in the eigenvalues. Figure 2 shows that the first 3 components for size and labour productivity explain at least 60% of the total density time variation, therefore we proceed to analyze their time pattern next.

3.4 Estimated Dynamic Strength Components

Figure 3 contains the plots of the three principal components when analyzing the dynamic properties of the variable distributions. The time evolution of $\widehat{\theta}_{tr}$ is compared against $\widehat{\theta}_{1r}$, so that the estimated $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ for $r = 1, \ldots, 3$ are plotted. For each variable, we highlight the movements of the first three principal components, as suggested in Figure 2.

The first row contains the leverage $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$. All the components of the leverage distribution exhibit a high degree of persistence. The movement of $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ tracks the movement of $\widehat{\theta}_{tr}$ as $\widehat{\theta}_{1r}$ is constant. For the 1985 cohort, the first component has a steep drop, slight increase, and then flattens with a permanent negative difference between $\widehat{\theta}_{t,1}$ and $\widehat{\theta}_{1,1}$. The second and third components show some variability at the beginning of the time sample but then have a similar time pattern of falling below, at a constant rate, the corresponding component value at age one. The 1989 cohort shows the same pattern in the first component. However, the second and third components, although displaying some variability, do not show a remarked divergence from $\widehat{\theta}_{1r}$ as $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ is close to zero. For both cohorts, there are persistent changes occurring to their respective leverage distribution. Both cohorts experience variability in component values over the first five years, while after age five (1990 for the 1985 cohort and 1994 for the 1989 cohort) the value of $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ remains constant for all three components.

The results suggest a cohort's leverage distribution initially experiences turbulent movements but stabilizes relatively shortly after entry.

The second row displays these changes in estimated principal components when analyzing the logarithm of sales distribution. For 1985 cohort, each component is transitory with at least one incidence of a short-lived deviations as indicated by the spikes. The 1989 cohort is different in that the second component difference has a permanent and persistent decrease. For the 1985 cohort, the size distribution does appear to be somewhat stable after age 3 (1988), but spikes in $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ are observed at age eight (1993) for the second component and at age ten (1995) for the third component. For the 1989 cohort, the size distribution appears to stabilize after age five (1994) as the value of $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ remains constant for all three components. As will leverage, the size distribution appears to move towards an equilibrium distribution with most of the movement occurring in the first five years of an entry cohort's life.

Finally, for labour productivity the deviations in the principal components for labour productivity distribution are for the most part transitory. Unlike the other variables, stability occurs at early ages followed by periods of instability. The 1985 cohort shows more transitory and volatility in the differenced components. The volatility happens around 1989 and in the 1990s. The 1989 cohorts are for the most part stable until the end of the time period. Both the 1985 and 1989 cohort suffer a spike in 1994. After 1994, the values of $\widehat{\theta}_{tr} - \widehat{\theta}_{1r}$ remain stable at close to zero for the 1985 cohort, while the component values continue to move around for the 1989 cohort. Age effects are a possible explanation for the observed movements. Aggregate effects provide alternative explanations. A recession occurred in Canada between 1990 and 1992. Further, the Canada-US free trade agreement (CUSFTA) was enacted in 1989, while the North American free trade agreement (NAFTA) came into effect in 1994. These events probably had an impact, at least temporarily, on individual firm's productivity.

4 Empirical Evidence from Static Distributions

Having analyzed the joint dynamic behaviour of the leverage, size and labour productivity distribution, and in view of results in Table 1 we proceed to look at the the source of the observed time changes for each distribution.

4.1 Finite Mixture Decompositions

The finite mixture models are used to understand the changes in heterogeneity over time and between cohorts of the outcome variables of interest. Our analysis of the 1985 and 1989 entry cohorts moves to checking for the presence of mixtures in the individual variable distributions at cross-sectional level. We estimate univariate finite mixtures using the methodology described in McLachlan and Peel (2000). There are $2 \times t_j$ decompositions of the distributions of each variable of interest requiring $2 \times t_j$ estimates of the K-class probability density functions, with t_j being the length of time for cohort j. In this case, if $f_k(y)$ is the k^{th} class probability density function and p_k denotes the proportion of the k^{th} class, then the finite mixture model with K components is defined as $f(y) = \sum_{k=1}^{K} p_k f_k(y)$. The proportion p_k can be interpreted as the prior probability of observing a sample from class k, with the property that is greater or equal to zero and that they sum to one ($p_k \geq 0$ and $\sum_{k=1}^{K} p_k = 1$). The objective is to estimate the parameters of the class probability densities and the proportions p_k of each class, while fixing an upper bound on the number of possible classes. The class of probabilities densities are assumed to have parametric components that can be estimated. We tested different parametric specifications and chose the ones that fit better the true distributions. BIC measures and tests of equality of distributions are used to evaluate which parametric specification fits better the true data.

The parameters of such mixtures can be estimated by maximum likelihood. For labour productivity and logarithm of sales, a mixture of normal distributions is estimated with the k^{th} probability density function given by $f_k(y|\theta) = \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left(\frac{-(y-\mu_k)^2}{2\sigma_k^2}\right)$. A mixture of lognormal distributions is used for the leverage with the k^{th} probability density function given by $f_k(y|\theta) = \frac{1}{y\sigma_k\sqrt{2\pi}} \exp\left(\frac{-(\log y-\mu_k)^2}{2\sigma_k^2}\right)$, where y is the variable of interest and the parameters of interest are $\theta = \{K, p_k, \mu_k, \sigma_k\}^\top$ with $k = 1, ..., K$ and K is the number of mixtures. The parameter p_k gives the weight on each mixture, $\sum_{k=1}^{K} p_k = 1$, and μ_k, σ_k provide the mean and standard deviation for each mixture respectively. All the parameters of interest with the exception of the number of types are estimated by the maximum likelihood. The number of types is selected using the AIC criteria, i.e. $\text{AIC}_k = -2\log l\left(\theta|y\right) + 2d_k$ is minimized with respect to k, where d_k is equal to the dimension of the model and acts as a correction term without which one would choose the model that maximizes the unconditional log-likelihood.

Tables 3 and 4 present the results of finite mixtures analysis for the distribution of log-sales for the 1985 and 1989 entry cohort, respectively. There are various similarities as well as differences between the two cohorts. At age one, the log-sales distribution for both cohorts is composed of three mixtures, while at age four and seven two mixtures are identified for both cohorts. At age one both cohorts have the highest share of sales associated to medium size firms, while for ages four and seven both cohorts have the highest share of sales associated with the smaller firms. The 1985 cohort presents an interesting feature at the age of four, when the 1989 cohort enters. It shows a concentration of the sales for the smaller firms (72 percent). In terms of mean sales by type, cohort 1985 has higher means than cohort 1989 at ages one and four. At age seven, when compared with 1989 cohort, cohort 1985 has a higher mean for the highest shares firms, but lower means for the other firms. The variances of the sales types are showing some variation between the cohorts at different ages of the firms. The difference is that cohort 1989 shows higher variation than cohort 1985 at ages one and seven while the reverse is true at age four.

Table 5 and 6 examine the presence of finite mixtures for the leverage distribution of the 1985 and 1989 entry cohorts, respectively. For this outcome variable, more differences are observed between the two cohorts than for the previous outcome variable. The same number of mixtures (three) are identified for ages one and four for both cohorts, with cohort 1989 keeping the same number of mixtures at age seven while cohort 1985 showing only two mixtures at that age. For cohort 1985 the low leverage and medium leverage firms are dominating the distribution at ages one and four, while for the cohort 1989 only at age seven the low and medium leverage are dominating the distribution, for the other ages the low and high leverage firms are dominating the distribution. In the same tone, cohort 1985 presents less concentration of types of firms than cohort 1989, which shows a higher concentration of firms with medium leverage. For cohort 1985 there is steady decrease in leverage over time, with very small changes between age four and seven. In contrast, there is a steady decrease in mean leverage for the 1989 cohort only for ages one and four and an increase in leverage during year seven. Finally, there is a higher variation in leverage for cohort 1989 than for cohort 1985 especially at age seven.

Finite mixture analysis for labour productivity is presented in tables 7 and 8 for the

1985 and 1989 cohorts, respectively. For both cohorts, there is a reduction in the number of mixture types. By age seven, there are only two types for both cohorts: low and high productivity. The difference in mean productivity between the high and low types is small between 0.4 and 0.7 with even smaller differences for cohort 1989. While the mean estimated labour productivity values for these two groups are almost equal at all ages, the estimated standard deviations are very different and provide the distinguishing factor between the estimated types.

4.1.1 Coefficient of Variation of Mixture Components

Figure 4 illustrates the results of the finite mixture components of all time periods for the 1985 and 1989 cohort. The information is collapsed into a measure of a weighted coefficient of variation: $CV_k = p_k\mu_k/\sigma_k$, where $k = 1, \ldots, K$, p_k is the probability of the mixture k, μ_k is the mean of the mixture k, and σ_k is the standard deviation of the mixture k. The figures have discontinuous plots since for some cases the mixture component is not present in certain years. The first row of the figure shows a downward trend in the first and second mixture component for leverage in both the 1985 and 1989 component. This confirms the density pictures which illustrates that the leverage distribution is shifting leftwards. The third component does not show a downward trend but rather volatility.

The firm size finite mixture components is in the second row. The first stark pattern is that the first component is always larger than the second component for both cohorts. There is a trend of the first component increasing while the second component is decreasing. This trend is indicative of the first component is getting larger while the second component is getting smaller. The difference between the two components is larger in the 1985 cohort relative to the 1989 cohort. Once again, this confirms the density estimation that shows that 1985 cohort is more heterogeneous.

The third row contains the labour productivity finite mixture components. The first component is larger than the second component for the most part in both cohorts. In the 1985 cohort there is marked decrease in the first component and increase in the second component in 1993. Part of it maybe due to an emergence of a third component which may decrease the weight in the first component and increase it for the second component. A similar case is shown for the 1989 cohort in year 1991 where the emergence of another component causes a large change in the first and second component.

4.2 Stochastic Dominance in the Presence of Finite Mixtures

To complete the static analysis we need to integrate the information obtained from finite mixture decompositions into a unified measure of comparison, which can be done using stochastic dominance testing. These tests are used to compare whether pair of distributions are equal and if not what it is the level of dominance, i.e. equality of two distributions (EoD), first-order stochastic dominance (FOSD) and second-order stochastic dominance (SOSD), as explained in Table 2.

Let $Y_{i,t;l}$ represent either leverage, size or debt-to-asset ratio of firm $i = 1, \ldots, n_t$, at age t, in cohort l, and let its distribution be $F_{t;l}(y) \equiv \Pr[Y_{i,t;l} \leq y]$. Analogous, let $Y_{i,s;j}$ represent the outcome variable of interest of firm $i = 1, \ldots, m_s$, at age s, in cohort j with its corresponding distribution $F_{s;j}(y) \equiv \Pr[Y_{i,s;j} \leq y]$. Furthermore, let $D_1^{[t;l]}(y) = F_{t;l}(y)$, and define higher orders for $D(*)$ as $D_o^{[t;l]}(y) = \int_0^y D_{o-1}^{[t;l]}(x)dx$. Note that we can also write $D_o^{[t;l]}(y) = 1/(o-1)! \int_0^y (y-x)^{o-1} dF_{t;l}(x)$. Table 2 provides the null hypothesis and the test

Table 2: Test of Distributions

	H_0	Test Statistic
EoD:	$F_{t;l}(y) = F_{s;j}(y)$	$\widehat{K} = \sqrt{\frac{n_t m_s}{n_t + m_s}}\ \sup_x \lvert \widehat{F}_{t;l}(y) - \widehat{F}_{s;j}(y) \rvert$
FOSD:	$F_{t;l}(y) \leq F_{s;j}(y)$	$\widehat{D} = \sqrt{\frac{n_t m_s}{n_t + m_s}} \sup_x [\widehat{F}_{t;l}(y) - \widehat{F}_{s;j}(y)]$
SOSD:	$D_2^{[t;l]}(y) \geq D_2^{[s;j]}(y)$	$\widehat{H} = \sqrt{\frac{n_t m_s}{n_t + m_s}} \sup_x [\widehat{D}_2^{[t;l]}(y) \geq \widehat{D}_2^{[s;j]}(y)]$

[a] Linton, Maasoumi, and Whang (2005) is used to test EoD and FOSD.
[b] Andrews (2000) and Dufour (2006) are used to test SOSD.

statistics for each case.

Linton, Maasoumi, and Whang (2005) is used to test EoD and FOSD, while Andrews (2000) and Dufour (2006) are used to test SOSD. However, since the above test statistics depend on nuisance parameters (the number of finite mixtures in each variable's distribution), their critical values are tabulated by the following parametric bootstrap procedure:

1. Sample n_t and n_s values, $(Y^*_{1,t;l}, \ldots, Y^*_{n_t,t;l})$ and $(Y^*_{1,s;j}, \ldots, Y^*_{n_s,s;j})$, from the estimated distributions $\int_0^y \widehat{f}(u)du = \int_0^y \sum_{k=1}^K \widehat{p}_k \widehat{f}_k(u)\,du$.

2. Using the bootstrap samples in step 1, calculate the empirical distribution functions $\widehat{F}^*_{t;l}(y)$ and $\widehat{F}^*_{s;j}(y)$, and calculate $\widehat{K}^*$, $\widehat{D}^*$, and $\widehat{H}^*$accordingly.

3. Repeat steps 1-2 B many times and define the critical value as the smallest value of y subject to at least $100(1-\alpha)\%$ of the obtained B values of $\widehat{K}^*$, $\widehat{D}^*$, and $\widehat{H}^*$ are at or below y.

4. Reject if $\widehat{K}^*$, $\widehat{D}^*$, or $\widehat{H}^*$ are greater than the critical value found in step 3.

4.2.1 Results of Stochastic Dominance Tests

The presentation of the tests refers to a) within a cohort comparisons. Analogous tests are used for b) between cohorts comparisons. Table 9 presents distribution comparisons over time within each cohort, i.e. $t \neq s$ and $l = j$. For the 1985 log-sales distribution, age one is SOSD dominated by ages four and seven, with ages four and seven stochastically equal. For the 1989 cohort, the age one is equal to age four and both are SOSD by age seven. The finding may suggest that cohort 1985 grew faster than cohort 1989.

The leverage story is very similar with the log-sales story but in opposite direction and with a higher magnitude. Therefore, for the leverage distribution, age one SOSD age four and seven for both cohorts, while the age four distribution does not look stochastically different than the age seven distribution. This finding indicates that for cohort 1985 we may have a convergence for the leverage distribution after age four. The convergence to an equilibrium leverage distribution appears to be longer for 1989 cohort as we find this cohort's age four leverage distribution SOSD its age seven distribution. Labour productivity does not show any changes in its distribution as we cannot reject equality of distributions across all ages for both cohorts.

Table 10 presents the comparison of the distributions between the entry cohorts, i.e. $t = s$ and $l \neq j$. For log-sales variable, cohort 1985 SOSD cohort 1989 at age four, while the two cohorts distributions at age one and seven are stochastically equal. In the leverage case the 1989 cohort SOSD leverage distribution for the 1985 cohort at all ages. The finding may suggest that pre free trade entry firms needed less leverage to survive than post free trade entry firms as they were more protected by differential tariffs. In contrast, for the labour productivity distribution we do not find any statistical difference between the cohorts at any ages.

5 Concluding Remarks

Our analysis looks at the evolution of debt-to-asset ratio, size and productivity distributions for two entry cohorts (1985 and 1989) of manufacturing firms. Functional principal component analysis provides a concise method to visualize how distributions evolve over time. Functional principal component analysis provides a method to capture dynamics and breaks a variable's distribution into dynamic principal components. The results show deviations in each entry cohort's debt-to-asset ratio distribution are persistent, while deviations are transitory for productivity and size distributions. A persistent component allows a variable's distribution to move around a steady state distribution. Transitory components capture changes in a variable's distribution over time, such as a movement towards an equilibrium distribution. The transitory components can account for deviations in a distribution occurring over time.

Static analysis makes pairwise comparisons of distributions using stochastic dominance tests and finite mixture analysis. The pairwise static distributional analysis confirms the findings of the dynamic functional principal component analysis. The distribution for the debt-to-asset ratio moves across time, while the distributions for productivity and size tend to be more stable over time. The functional principal component analysis and the static pairwise analysis provide effective methods to describe the evolution of a variable's distribution over time with minimal assumptions.

References

Andrews, D. W. K. (2000): "Inconsistency of the Bootstrap when a Parameter Is on the Boundary of the Parameter Space," *Econometrica*, 68(2), 399–406.

Angelini, P., and A. Generale (2008): "On the Evolution of Firm Size Distributions," *American Economic Review*, 98(1), 426–38.

Bai, J., and S. Ng (2002): "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.

Cabral, L. M., and J. Mata (2003): "On the evolution of the firm size distribution: Facts and theory," *American Economic Review*, 93(4), 1075–1090.

Dufour, J.-M. (2006): "Monte Carlo tests with nuisance parameters: A general approach to finite-sample inference and nonstandard asymptotics," *Journal of Econometrics*, 133(2), 443–477.

Gibrat, R. (1931): *Les Inegalités Economiqués.* Librairie du Recueil Sirey. Paris.

Huynh, K. P., and R. J. Petrunia (2008): "Age Effects, Leverage, and Firm Growth," mimeo.

Huynh, K. P., R. J. Petrunia, and M. C. Voia (forthcoming): "The Impact of Initial Financial State on Firm Duration Across Entry Cohorts," *Journal of Industrial Economics*, forthcoming.

Jovanovic, B. (1982): "Selection and the Evolution of Industry," *Econometrica*, 50, 649–670.

Kneip, A., and K. J. Utikal (2001): "Inference for Density Families Using Functional Principal Component Analysis," *Journal of the American Statistical Association*, 96(454), 519–542.

Linton, O., E. Maasoumi, and Y.-J. Whang (2005): "Consistent Testing for Stochastic Dominance under General Sampling Schemes," *Review of Economic Studies*, 72(3), 735–765.

Lucas, R. E. (1978): "On the Size Distribution of Business Firms," *Bell Journal of Economics*, 9(2), 508–523.

McLachlan, G., and D. Peel (2000): *Finite Mixture Models.* Wiley, New York.

Pesaran, M. H. (2006): "Estimation and Inference in Large Heterogeneous Panels with a Multifactor Error Structure," *Econometrica*, 74(4), 967–1012.

Table 3: Log of Size (Sales) Distribution: Cohort 1985

Year	Type	Share	μ	s.e.	σ	s.e.	logL	N
1986	I	0.610	5.861	0.079	1.085	0.131	-4371.73	2578
1986	II	0.247	6.415	0.151	2.048	0.135		
1986	III	0.143	5.373	0.113	0.501	0.216		
1989	I	0.717	6.206	0.052	1.017	0.067	-3377.69	1993
1989	II	0.283	7.048	0.155	1.828	0.110		
1992	I	0.697	6.242	0.087	1.094	0.116	-2804.305	1615
1992	II	0.303	7.005	0.205	1.813	0.159		

Table 4: Log of Size (Sales) Distribution: Cohort 1989

Year	Type	Share	μ	s.e.	σ	s.e.	logL	N
1990	I	0.506	5.93	0.217	1.07	0.159	-3793.08	2266
1990	II	0.291	5.232	0.129	0.681	0.193		
1990	III	0.203	6.175	0.161	2.222	0.162		
1993	I	0.536	5.771	0.086	0.978	0.143	-3049.33	1710
1993	II	0.464	6.522	0.133	1.808	0.106		
1996	I	0.838	6.138	0.121	1.276	0.077	-2558.86	1402
1996	II	0.162	8.098	1.179	1.657	0.349		

Note: The parameters of the finite mixtures: number of types (k), mean of type k (μ_k), and standard deviation of type k (σ_k) were obtained by maximizing the log-likelihood associated to the following density function:

$$f_{\text{log-sales}}(y|\theta) \quad = \quad \sum_{k=1}^{K} p_k \frac{1}{\sigma_k\sqrt{2\pi}} \exp\left(\frac{-(y-\mu_k)^2}{2\sigma_k^2}\right).$$

The number of types were chosen according to the number for which the likelihood is the highest and the other parameters of interest are stable.

Table 5: Leverage Distribution: Cohort 1985

Year	Type	Share	μ	s.e.	σ	s.e.	logL	N
1986	I	0.514	0.822	0.025	0.365	0.062	-1596.73	2578
1986	II	0.354	0.931	0.013	0.149	0.051		
1986	III	0.132	1.018	0.080	1.584	0.103		
1989	I	0.391	0.830	0.024	0.181	0.068	-1687.58	1993
1989	II	0.378	0.697	0.061	0.375	0.177		
1989	III	0.231	0.716	0.115	0.996	0.101		
1992	I	0.581	0.696	0.051	0.685	0.065	-1606.77	1615
1992	II	0.419	0.812	0.017	0.212	0.061		

Table 6: Leverage Distribution: Cohort 1989

Year	Type	Share	μ	s.e.	σ	s.e.	logL	N
1990	I	0.718	0.966	0.013	0.239	0.042		
1990	II	0.224	1.167	0.057	1.746	0.076	-1508.82	2266
1990	III	0.058	0.514	0.141	0.131	0.105		
1993	I	0.618	0.911	0.028	0.239	0.060		
1993	II	0.250	1.191	0.065	1.886	0.085	-1441.67	1710
1993	III	0.132	0.459	0.176	0.149	0.125		
1996	I	0.519	0.819	0.019	0.240	0.069		
1996	II	0.389	0.699	0.086	0.597	0.164	-860.71	1402
1996	III	0.092	1.736	0.762	4.959	0.301		

Note: The parameters of the finite mixtures: number of types (k), mean of type k (μ_k), and standard deviation of type k (σ_k) were obtained by maximizing the log-likelihood associated to the following density function:

$$f_{\text{leverage}}(y|\theta) \quad = \quad \sum_{k=1}^{K} p_k \frac{1}{y\sigma_k\sqrt{2\pi}} \exp\left(\frac{-(\log y - \mu_k)^2}{2\sigma_k^2}\right).$$

The number of types were chosen according to the number for which the likelihood is the highest and the other parameters of interest are stable.

Table 7: Labour Productivity Distribution: Cohort 1985

Year	Type	Share	μ	s.e.	σ	s.e.	logL	N
1986	I	0.702	4.342	0.035	0.543	0.071	-3084.8	2578
1986	II	0.224	5.052	0.418	0.843	0.204		
1986	III	0.074	4.299	0.221	2.111	0.178		
1989	I	0.755	4.428	0.018	0.473	0.043	-2021.03	1993
1989	II	0.245	4.852	0.071	1.125	0.084		
1992	I	0.814	4.492	0.018	0.493	0.039	-1643.81	1615
1992	II	0.186	4.9	0.093	1.286	0.111		

Table 8: Labour Productivity Distribution: Cohort 1989

Year	Type	Share	μ	s.e.	σ	s.e.	logL	N
1990	I	0.531	4.432	0.037	0.455	0.077	-2744.68	2266
1990	II	0.376	4.887	0.105	0.857	0.137		
1990	III	0.093	4.569	0.201	2.081	0.201		
1993	I	0.780	4.630	0.020	0.545	0.043	-2001.57	1710
1993	II	0.220	4.675	0.088	1.491	0.110		
1996	I	0.715	4.594	0.024	0.481	0.05	-1472.24	1402
1996	II	0.285	4.976	0.074	1.091	0.091		

Note: The parameters of the finite mixtures: number of types (k), mean of type k (μ_k), and standard deviation of type k (σ_k) were obtained by maximizing the log-likelihood associated to the following density function:

$$f_{\text{labour productivity}}(y|\theta) \quad = \quad \sum_{k=1}^{K} p_k \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left(\frac{-(y-\mu_k)^2}{2\sigma_k^2}\right).$$

The number of types were chosen according to the number for which the likelihood is the highest and the other parameters of interest are stable.

Table 9: Stochastic Dominance: Within Cohorts

1985	Sales: Age 1	Sales: Age 4	Sales: Age 7	Leverage: Age 1	Leverage: Age 4	Leverage: Age 7	Labour Productivity: Age 1	Labour Productivity: Age 4	Labour Productivity: Age 7
Age 1	eq	-2	-2	eq	2	2	eq	eq	eq
Age 4	.	eq	eq	.	eq	eq	.	eq	eq
Age 7	.	.	eq	.	.	eq	.	.	eq

1989	Sales: Age 1	Sales: Age 4	Sales: Age 7	Leverage: Age 1	Leverage: Age 4	Leverage: Age 7	Labour Productivity: Age 1	Labour Productivity: Age 4	Labour Productivity: Age 7
Age 1	eq	eq	-2	eq	2	2	eq	eq	eq
Age 4	.	eq	-2	.	eq	2	.	eq	eq
Age 7	.	.	eq	.	.	eq	.	.	eq

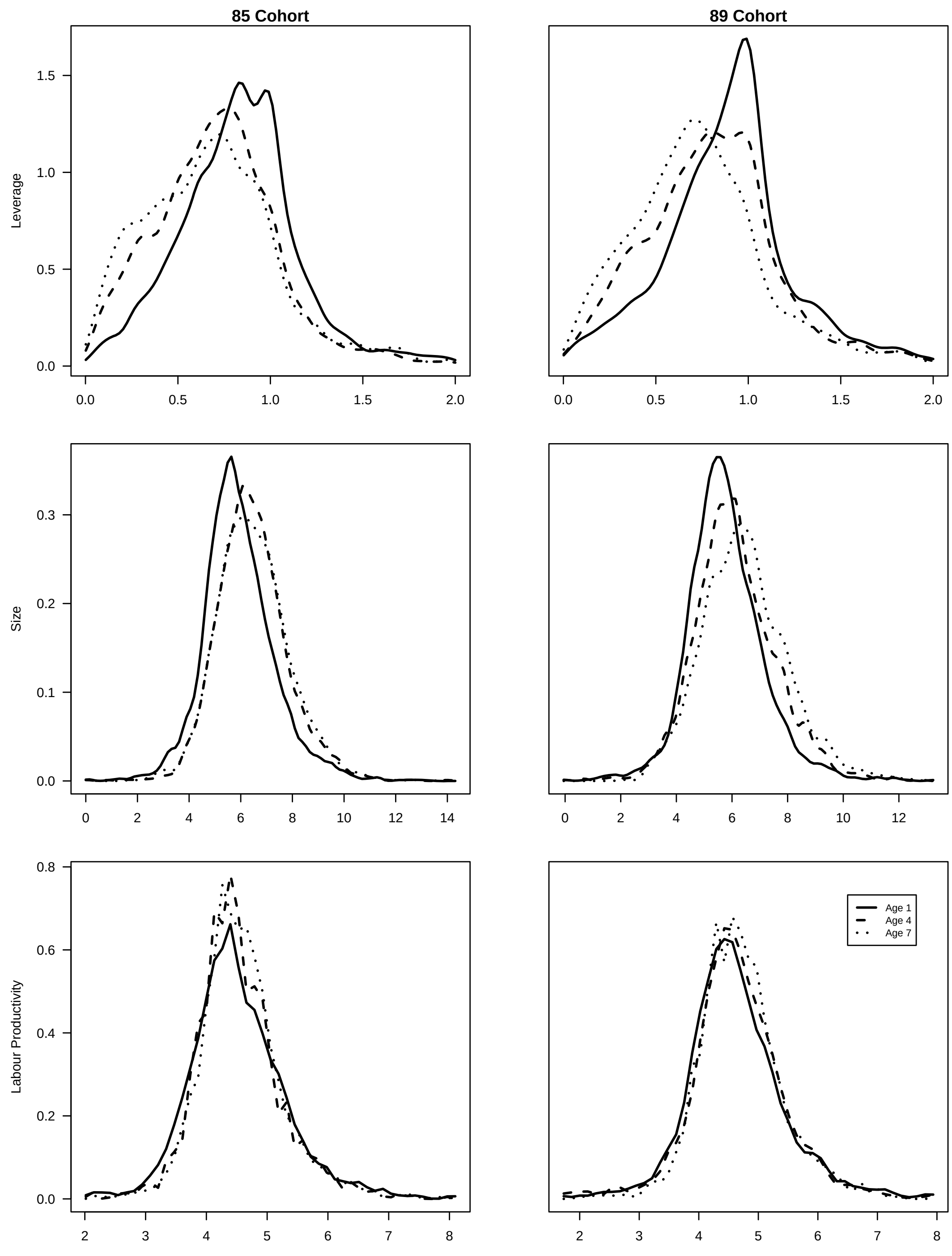
Note: The results are read in terms of row versus columns. (eq) stands for equality of distributions, (2) stands for second order stochastic dominance and a negative number indicates that the column dominates the row, here (-2) stands for second order stochastic dominated distribution. For example, for the 1985 cohort, line one (columns 2, 3 and 4 - for leverage) reads that Age 1 SOSD both Age 4 and Age 7, while line two (columns 3 and 4 - for leverage) reads that Age 4 is stochastically equal to Age 7 distribution. All results are significant at 5%.

Table 10: Stochastic Dominance: Between Cohort 1985 & 1989

	Sales 1985 vs. 1989	Leverage 1985 vs. 1989	Labour Productivity 1985 vs. 1989
Age 1	eq	-2	eq
Age 4	2	-2	eq
Age 7	eq	-2	eq

Note: The results are read in terms of row versus columns. For example column 2 (for leverage) reads that at each age cohort 1989 SOSD cohort 1985. All results are significant at 5%.

Figure 1: Distributions

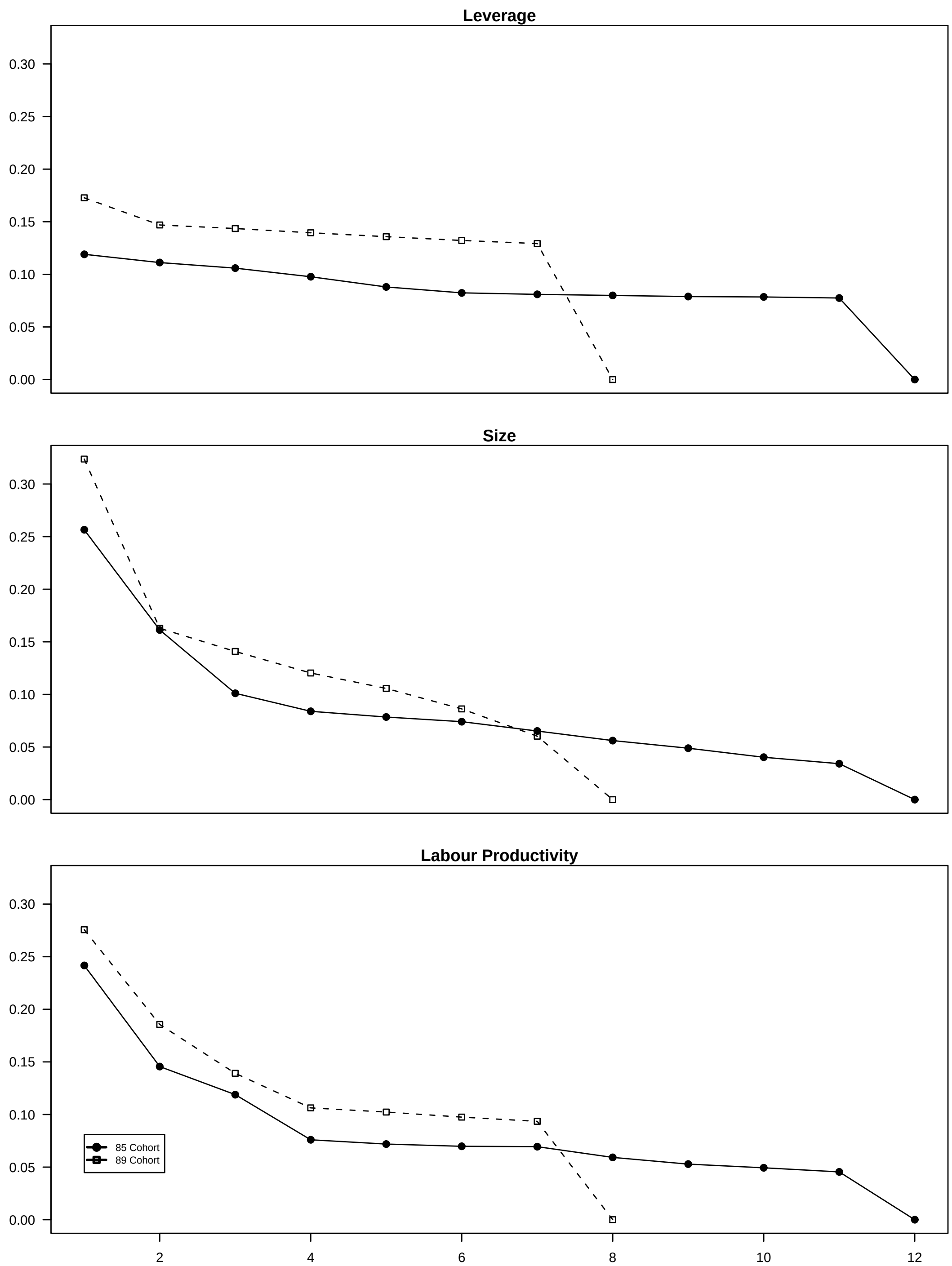


[a] Cross-validated bandwidths are 0.2598, 0.3214, and 0.3585 for firms aged 1, 4 and 7-years old respectively in the 1985 cohort (First column).

[b] Cross-validated bandwidths are 0.2642, 0.2707, and 0.3114 for firms aged 1, 4 and 7-years old respectively in the 1989 cohort (Second column).

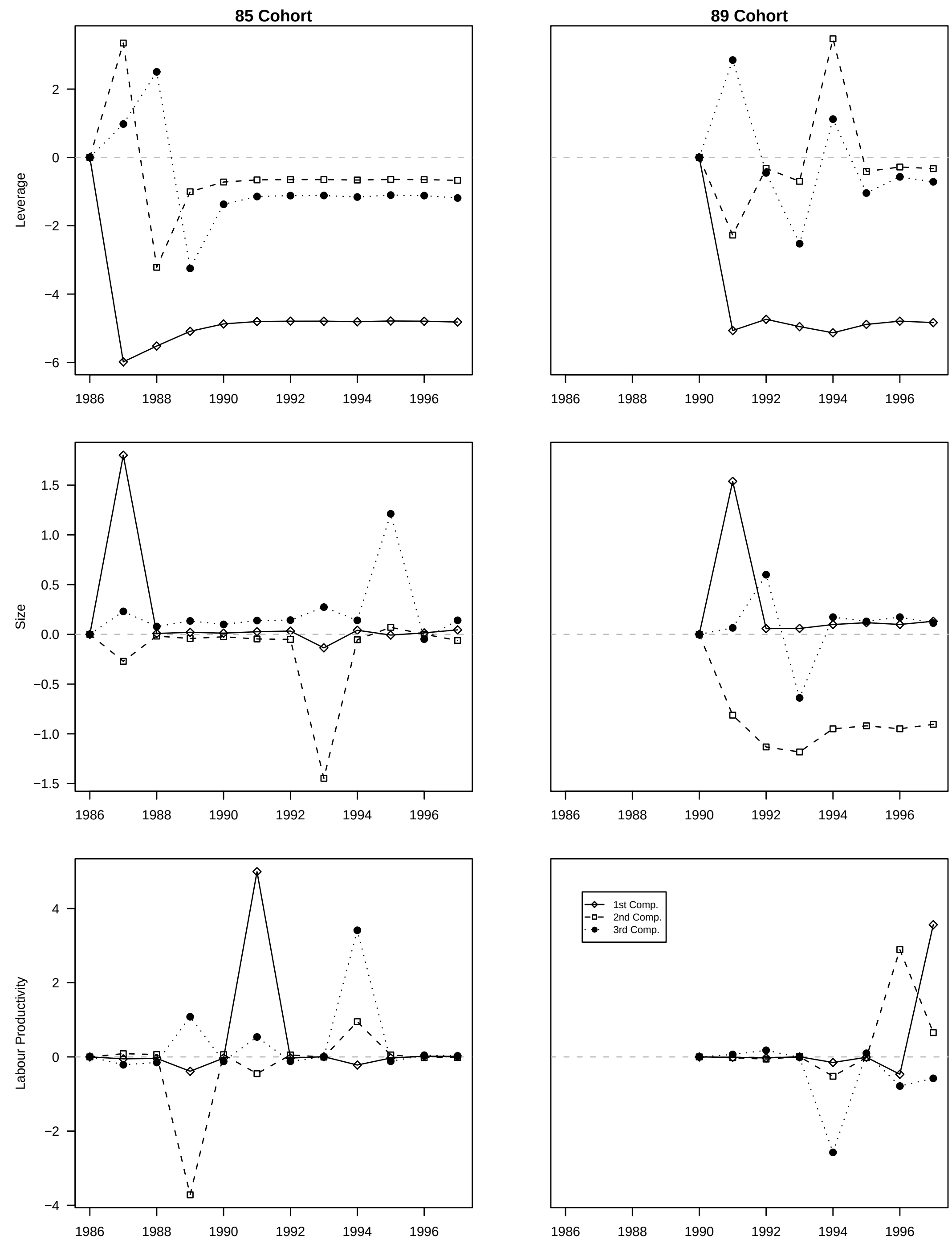
[c] Gaussian kernels are used throughout.

Figure 2: Dynamic Scree Plot



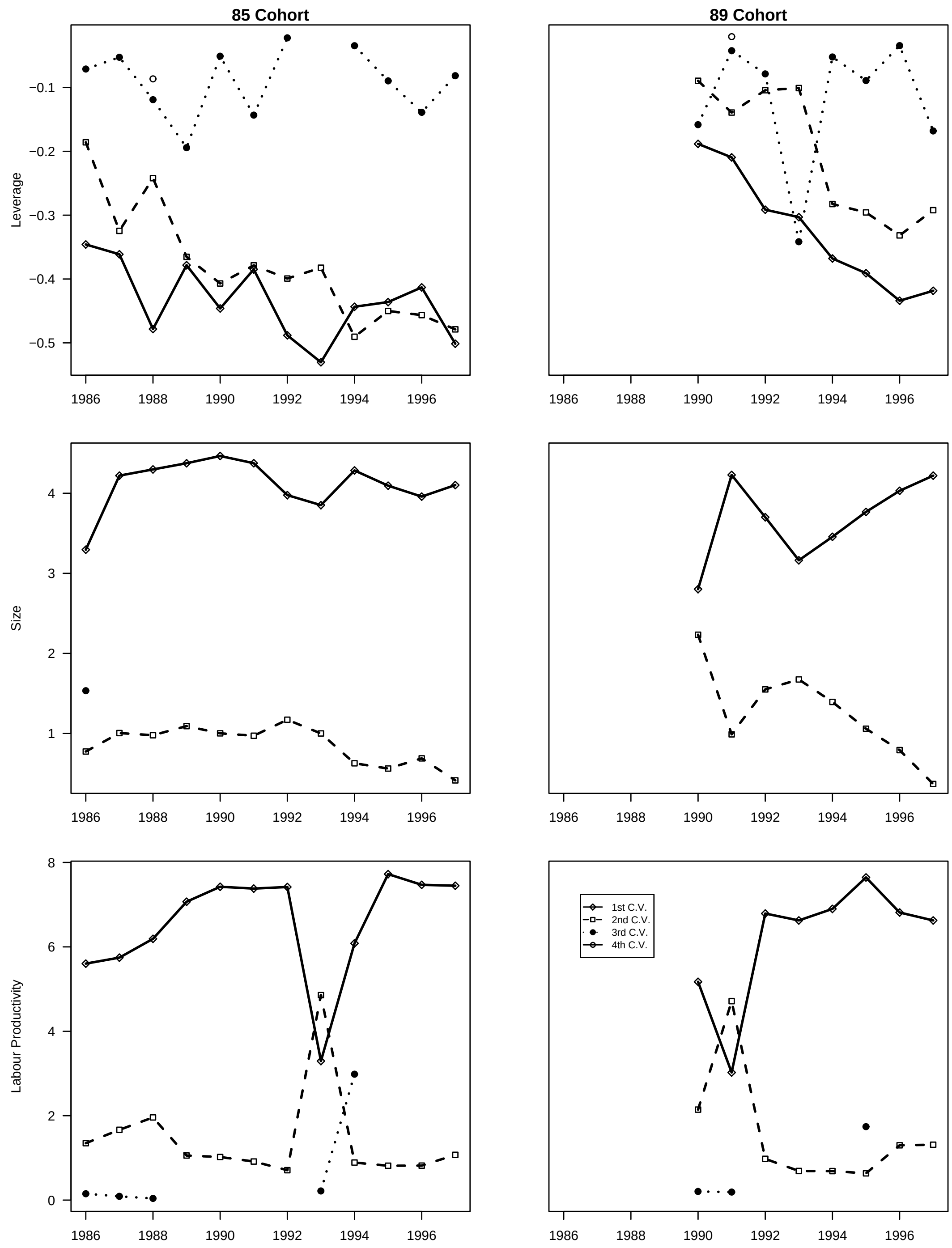
Note: Each plot displays $\widehat{\lambda}_r / \sum_{r=1}^{T} \widehat{\lambda}_r$ for the 1985 and 1989 cohorts.

Figure 3: Estimated Dynamic Strength Components



Note: Each figure plots $\widehat{\theta}_{tr} - \widehat{\theta}_{t_0 r}$ for $r = 1, 2, 3$.

Figure 4: Coefficient of Variation



Note: Each plot displays $CV_k = p_k \mu_k / \sigma_k$ where $k = 1, \ldots, K$, p_k is the probability of the mixture k, μ_k is the mean of the mixture k, and σ_k is the standard deviation of the mixture k.