Cardiff Economics Working Papers

Working Paper No. E2022/15

The Confidence Interval of Cross-Sectional Distribution of Durations

Huw Dixon and Maoshan Tian

August 2022
(Updated January 2023)

ISSN 1749-6010

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

This working paper is produced for discussion purpose only. These working papers are expected to be published in due course, in revised form, and should not be quoted or cited without the author's written permission.
Cardiff Economics Working Papers are available online from:
http://econpapers.repec.org/paper/cdfwpaper/ and
business.cardiff.ac.uk/research/academic-sections/economics/working-papers
Enquiries: EconWP@cardiff.ac.uk

The Confidence Interval of the Cross-Sectional Distribution of Durations *

Huw Dixon[1] and Maoshan Tian[2‡]

Abstract

Tian and Dixon (2022) derived the variance of the estimator of cross-sectional distribution of durations (CSD). In this paper, we apply both Fieller's method and the Delta method to derive confidence interval of CSD using this variance formula. (CSD) is a new estimator derived by Dixon (2012). It can be applied in general Taylor model (GTE) by Dixon and Bihan (2012) and hospital waiting times by Dixon and Siciliani (2009). We use Monte Carlo simulations to evaluate the empirical size of Fieller's method and delta method among different sample sizes. The empirical results show that both Fieller's method and the delta method are valid in terms of estimating the confidence interval of CSD. Finally, we use both methods for real data set: the UK CPI micro-price data. Depending on the application, we see that both methods provide reasonable CIs for CSD estimators.

JEL Codes: C10, C15, E50

Keywords: Fieller's Method, Delta Method, Confidence Interval

*We are grateful for very helpful comments from Patrick Minford, Kul Luintel, Walter Distaso and seminar participants at Cardiff University and also from participants at the 2018 Econometric Society Australasia Meeting.

†1H. Dixon is a Professor at Cardiff Business School, University of Cardiff and Lead Researcher in Economic Measurement at the National Institute of Economic and Social Research (London) dixonh@cardiff.ac.uk

‡2Dr M. Tian is an assistant professor at School of Economics and Management, Chongqing University of Posts and Telecommunications. tianms@cqupt.edu.cn

1 Introduction

The non-parametric cross-sectional distribution of durations (CSD) is derived and introduced in Dixon (2012). It is a new estimator which has been applied to CPI micro-price data (see Dixon and Bihan (2012) which uses the French data, and Dixon and Tian (2017) which uses the UK data) and hospital waiting time data (see Dixon and Siciliani (2009)). In this paper we investigate the confidence intervals of CSD in order to determine the accuracy of the estimator. We consider two ways of doing this, Fieller's method and the delta method. CSD is a ratio estimator, and as such is particularly appropriate for the Fieller method. Fieller (1932) investigated and derived the general cumulative distribution formula for the ratio distribution $w = \frac{x}{y}$, where both x and y follow the normal distribution and are correlated with each other. Fieller (1954) focused on the distribution of the ratio w where x and y were independent from each other. In Fieller's method, the ratio variable is transformed into a linear function. The confidence interval of the ratio variable can be obtained by solving out the quadratic roots the linear function. The alternative and more general method is the delta method which can be employed even when the distribution of the ratio variable is unknown. Franz (2007) gave a very good literature review of Fieller's method and its inferences and development, and for applications see Beyene and Moineddin (2005).

Another related approach that we do not follow here is to derive the probability density function (*pdf*) of the estimator (SeeMarsaglia (1965),Cedilnik et al. (2004),Cedilnik et al. (2006)). In the case of the CSD, this is difficult, since the expression of the CSD is relatively complicated and the *pdf* of CSD may not be a familiar distribution.[1] Therefore, we will focus on the confidence interval rather than the *pdf* of CSD. Another approach used by Carvalho et al. (2020) is to estimate the CSD using Bayesian methods in the context of a DSGE macroeconomic model, to infer CSD from the macroeconomic behaviour of the economic data. This is very different from the non-parametric approach used here and is clearly highly dependent on the model of the economic system employed.

Tian and Dixon (2022) derived the variance formulae for CSD and two other related non-parametric estimators (the distribution of durations or unconditional hazard DD, the cross-sectional distribution of ages CSA) and

[1]The distributions of both the survival and hazard functions affect CSD.

the covariance of the Kaplan-Meier (KM) estimators. Using these results, we can apply both the Fieller and delta methods to construct the confidence intervals for the CSD. There are some studies which focus on the comparison of the Fieller's method with other numerical methods. As Polsky et al. (1997) and Briggs et al. (1999) showed, Fieller's method is indeed suitable for constructing the confidence interval of ratio variables. Fan and Zhou (2007) suggested that the Fieller method, the standard bootstrap method and the bootstrap percentile method all provided accurate confidence intervals of ratio variables even when the numerator and denominator follow different distributions. Wang and Zhao (2008) suggested that the bootstrap Fieller method provided more accurate confidence interval of the ratio variable. In Bebu et al. (2016)'s simulation studies, the Fieller method provided the most accurate confidence interval for the ratio variable than the other methods. See also Cox (1990) and Gardiner et al. (2001).

In section 2 we review some basics of the non-parametric estimators of the survival and hazard functions, CSD and the related distributions DD and CSA. In section 3 we derive the CIs using the Fieller method: Theorem 1 states the test statistic and Theorem 2 the corresponding CIs for CSD (Corollary 1 and 2 do the same for CSA). Lemmas 1-3 derive the CIs for CSD, CSA and DD using the delta method. In section 4 we undertake a Monte Carlo analysis of the empirical size of the Fieller and delta methods for different sample sizes and significance levels. In section 5 we apply the method to empirical data, the UK CPI price-quote data from 1998 to 2017.

2 The KM Estimators of the Survival Function

In this section, we provide a brief summary of the main points about estimating the discrete time survival and hazard functions that we require for this paper. Kaplan and Meier (1958) provided a non-parametric estimator for the survival function, the Kaplan-Meier (KM) estimators of the survival probabilities $S_i \in [0, 1]$ for $i = 0, 1, 2, ..., F$, where F is the maximum duration observed in the data set. The survival probability gives the probability that the the agent remains in the same state for (strictly) more than i periods. The actual period the agent remains in the same state is called a spell. In the empirical application we use in this paper, the agent is a price-setter, the state is a particular price and a price-spell is an event where the same price is set by the same price-setter for a number of periods (the completed duration of the price-spell). In this paper we will only consider the case where we observe all spells in their entirety (all spells are uncensored).

If we look across the entire data set, we can count the number of spells that last at least k periods as N_k, and the number of failures in the k-th period as D_k. Failure in this context means that the spell has come to an end (for example, the price has changed). N_0=N is the total number of price spells in the initial period. The Kaplan-Meier estimator $\hat{S}_i$ is:

$$\hat{S}_i = \prod_{k=1}^{i} \frac{N_k - D_k}{N_k} \tag{1}$$

$\hat{S}_i$ can be defined as the proportion of spells surviving for (strictly) longer than i periods. We can set $\hat{S}_0 = 1$, since all spells last longer than zero periods and $\hat{S}_F$=0 since no spell lasts more than F periods.[2] Hence there are $F - 1$ survival probabilities to be estimated from the data.

The hazard function h_i is a conditional probability, estimated as the proportion of failures amongst spells that have lasted i periods:

$$\hat{h}_i = \frac{D_i}{N_i} \tag{2}$$

We assume $D_0 = 0$ and $\hat{h}_0 = 0$ because all spells last at least 0 periods, and $\hat{h}_F = 1$ since F is the longest spell observed. There remain $F - 1$ hazards to be estimated. Clearly, the estimator of the hazard function can be transformed into the KM estimator and vice versa. There is thus a one-to-one mapping between the estimated hazard function and survival function. Note that equation (2) is also the maximum likelihood estimator of the hazard function.[3] Therefore, the KM estimator is derived from the maximum likelihood estimator of the hazard function.

Finally, we define two additional variables we will be using. First, the sum of the estimated survival probabilities $\hat{S} = \sum_{k=0}^{F} \hat{S}_k$, and secondly $\bar{h}$ as the reciprocal of this sum.

$$\bar{h} = \frac{1}{\sum_{k=0}^{F} \hat{S}_k} = \frac{1}{\hat{S}} \tag{3}$$

[2] We are only able to assume no spells last longer than F because spells are uncensored. For example, if there were a right-censored spell that lasted F periods, we could have $\hat{S}_F$¿0

[3] See the appendix of Tian and Dixon (2022) for a formal derivation.

The variable $\bar{h}$ is the average proportion of agents that fail each period, In effect, this is the same as what engineers call FIT (Failures in Time). In our example of price-setting, this is the average number of firms which change prices per month. The variable S is the average duration of spells.

2.1 The Cross-sectional Distribution of Durations CSD

The main focus of this paper is the CSD. As a preliminary, we can define the distribution of durations (DD), which is also known as the unconditional hazard function. It gives the probabilities that a duration will last i periods where $i = 1, 2, ..., F$. The estimator of the distribution of durations can be written either as:

$$\hat{a_i^d} = \hat{S}_{i-1}\hat{h}_i, \tag{4}$$

or equivalently

$$\hat{a_i^d} = \hat{S}_{i-1} - \hat{S}_i. \tag{5}$$

The cross-sectional distribution of (completed) durations CSD gives the probabilities that spells observed at a point in time will last for i periods. The estimator of CSD can be written as:

$$\hat{a}_i = \frac{i\hat{S}_{i-1}\hat{h}_i}{\hat{S}} \tag{6}$$

Closely related to CSD is the cross-sectional distribution of ages (incomplete durations), CSA. The estimator for CSA is:

$$\hat{a_i^A} = \frac{\hat{S}_{i-1}}{\hat{S}} \tag{7}$$

The definitions and relationships between the three distributions and their estimators are described in detail in Tian and Dixon (2022).

2.2 Properties of estimators

In this section, we will review some earlier results and make a couple of observations about the estimates of the survival functions, the sum of survival probabilities and the variances of the estimators.

Breslow and Crowley (1974) showed that for the estimated survival function $\hat{S}_i$ with $i = 0, 1, 2, ..., F$, the vector $\hat{V} = (\hat{S}_0, \hat{S}_1, ..., \hat{S}_F)$ follows the asymptotic multivariate normal distribution: $\sqrt{n}(\hat{V} - V) \overset{d.}{\sim} MN(0, \Sigma)$. Where Σ is the variance-covariance matrix for the vector $\hat{V}$, and the vector $E(V)$ is the mean value of each element in vector $\hat{V}$.

Turning to the sum of the survival probability estimates, since $\hat{S}_0 = 1$ and $\hat{S}_F = 0$[4]:

$$\hat{S} = \sum_{k=0}^{F} \hat{S}_k = 1 + \sum_{k=1}^{F} \hat{S}_k \tag{8}$$

This follows the asymptotic normal distribution $\hat{S} \overset{d.}{\sim} N(\mu_{\hat{S}}, \sigma^2_{\hat{S}})$, where $\mu_{\hat{S}}$ is $E(\hat{S})$ and $\sigma^2_{\hat{S}}$ is the variance of $\hat{S}$.

Note that each $\hat{S}_i$ follows the normal distribution asymptotically and hence $\hat{S}$ also. Together, $\hat{S}_i$ and $\hat{S}$ follow the asymptotic multivariate normal distribution:

Observation 1 :

$$\sqrt{N} \begin{bmatrix} \hat{S}_i - \mu_{\hat{S}_i} \\ \hat{S} - \mu_{\hat{S}} \end{bmatrix} \overset{d.}{\sim} MN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\hat{S}_i} & \sigma_{\hat{S}_i,\hat{S}} \\ \sigma_{\hat{S}_i,\hat{S}} & \sigma^2_{\hat{S}_i} \end{bmatrix} \right)$$

Where N is the adjusted sample size. $\mu_{\hat{S}_i}$ is the mean value of $\hat{S}_i$, $\mu_{\hat{S}} = 1 + \mu_{\hat{S}_1} + \mu_{\hat{S}_2} + \ldots\ldots + \mu_{\hat{S}_F}$;

$Var(\hat{S}_i) = \sigma^2_{\hat{S}_i}$; $Var(\hat{S}) = \sigma^2_{\hat{S}}$;

$Var(\hat{S}) = \sigma^2 = \sum_{k=1}^{F} \sigma^2_{\hat{S}_k} + 2\sum_{k=1}^{F}\sum_{j=1}^{F} \sigma^2_{\hat{S}_k,\hat{S}_j}$ for $k \neq j$;

$\sigma_{\hat{S}_i,\hat{S}}$ is the covariance of $\hat{S}_i$ and $\hat{S}$; $\sigma_{\hat{S}_i,\hat{S}} = Cov(\hat{S}_i, \hat{S}) = \sum_{k=1}^{F} Cov(\hat{S}_i, \hat{S}_k)$. The variance and the covariance formula of the survival function are found from equation (9) and (10) below.

The Greenwood formula Greenwood (1926) for the variance of the KM estimator can be written as:

$$\widehat{Var}(\hat{S}_i) = \hat{S}_i^2 [\sum_{k=1}^{i} \frac{D_k}{N_k(N_k - D_k)}] \tag{9}$$

[4]If there are censored observations including in the data set, $\hat{S}_F$ may not equal to zero. Therefore, we still remain $\hat{S}_F$ in our formula at this point.

For our analysis, it is crucial to investigate the covariance among the survival function estimates. It is clear that the survival estimate $\hat{S}_i$ is correlated with $\hat{S}_j$ for $0 < i, j < F$. Breslow and Crowley (1974) investigated the large sample properties of the hazard function estimator (2) and the survival function. They found that the off-diagonal variance-covariance matrix of the hazard functions are all equal to zero. If the hazard function $\hat{h}_i$ with $i = 1, 2, ..., F$ are collected by the vector $\hat{h}$, the joint distribution of $\hat{h}$ follows the Gaussian distribution asymptotically. They also show that the vector of survival function $V = (\hat{S}_1, ...\hat{S}_F)$ converges weakly to the Gaussian process. They derived the asymptotic covariance of the survival functions between different periods. Tsai et al. (1987) wrote a literature review to discuss the covariance properties of KM estimators. Tian and Dixon (2022) applied the Taylor expansion to derive the covariance of the KM estimators as follows (discrete time):

$$\widehat{Cov}(\hat{S}_i, \hat{S}_j) = \hat{S}_i \hat{S}_j [\sum_{k=1}^{i} \frac{D_k}{N_k(N_k - D_k)}] \; for \; i < j \tag{10}$$

Recall the estimators CSD and DD defined in (6) and (4), and that $\hat{S}_{i-1}\hat{h}_i = \hat{S}_{i-1} - \hat{S}_i$. From Observation 1, $\hat{S}_i$ and the summation of the survival function $\hat{S}$ follow the multivariate normal distribution asymptotically. In addition, it follows that $\hat{S}_{i-1}\hat{h}_i$ and $\hat{S}$ also follow the multivariate normal distribution asymptotically:

Observation 2 :

$$\sqrt{N} \begin{bmatrix} \hat{S}_{i-1}\hat{h}_i - \mu_{\hat{S}_{i-1}\hat{h}_i} \\ \hat{S} - \mu_{\hat{S}} \end{bmatrix} \overset{d.}{\sim} MN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma^2_{\hat{S}_{i-1}\hat{h}_i} & \sigma_{\hat{S}_{i-1}h_i,\hat{S}} \\ \sigma_{\hat{S}_{i-1}\hat{h}_i,\hat{S}} & \sigma^2_{\hat{S}} \end{bmatrix} \right)$$

Where N is adjusted sample size. $\sigma^2_{\hat{S}_{i-1}\hat{h}_i} = Var(\hat{S}_{i-1}\hat{h}_i)$; $\sigma_{\hat{S}_{i-1}\hat{h}_i,\hat{S}}$ is the covariance between $\hat{S}_{i-1}\hat{h}_i$ and $\hat{S}$; $Cov(i\hat{S}_{i-1}\hat{h}_i, \hat{S}) = i\sigma_{\hat{S}_{i-1}\hat{h}_i,\hat{S}} = i[\sum_{k=1}^{F} Cov(\hat{S}_{i-1}, \hat{S}_k) - \sum_{k=1}^{F} Cov(\hat{S}_i, \hat{S}_k)]$.

The covariance between $i\hat{S}_{i-1}\hat{h}_i$ and $\hat{S}_j$ can be derived as:

$$Cov(i\hat{S}_{i-1}\hat{h}_i, \hat{S}_j) = iCov(\hat{S}_{i-1} - \hat{S}_i, \hat{S}_j) = i[Cov(\hat{S}_{i-1}, \hat{S}_j) - Cov(\hat{S}_i, \hat{S}_j)]$$

In other words, the $Cov(i\hat{S}_{i-1}\hat{h}_i, \hat{S}_j)$ can be transformed to the covariance of $Cov(i\hat{S}_{i-1}\hat{h}_i, \hat{S})$:

$$\begin{aligned} Cov(i\hat{S}_{i-1}\hat{h}_i, \sum_{k=1}^{F}\hat{S}_k) &= i[Cov(\hat{S}_{i-1}, \sum_{k=1}^{F}\hat{S}_k) - Cov(\hat{S}_i, \sum_{k=1}^{F}\hat{S}_k)] \\ &= i[\sum_{k=1}^{F}Cov(\hat{S}_{i-1}, \hat{S}_k) - \sum_{k=1}^{F}Cov(\hat{S}_i, \hat{S}_k)] \end{aligned}$$

Note that $Cov(\hat{S}_i, \hat{S}) = \sigma_{\hat{S}_i,\hat{S}} = \sum_{k=1}^{F} Cov(\hat{S}_i, \hat{S}_k)$. Using equation (10), we can calculate $Cov(\hat{S}_i, \hat{S})$.

The variance of CSD is derived by Tian and Dixon (2022) which can be written as:

$$\widehat{Var}(\hat{a}_i) = i^2\frac{\widehat{Var}(\hat{S}_{i-1}\hat{h}_i)}{\hat{S}^2} + i^2\frac{\hat{S}_{i-1}^2\hat{h}_i^2\widehat{Var}(\hat{S})}{\hat{S}^4} - 2i^2\frac{\hat{S}_{i-1}\hat{h}_i\widehat{Cov}(\hat{S}_{i-1}\hat{h}_i, \hat{S})}{\hat{S}^3} \quad (11)$$

After reviewing those properties, Fieller's method can now be applied to construct the confidence interval.

3 Confidence Intervals for the Cross-Sectional Distributions

The Fieller method is designed to calculate the confidence intervals (CI) for ratio estimators such as CSD and CSA, which we derive in Theorem (CSD) and its corollary (CSA). The delta method can be also be used for non-ratio estimators such as DD, so we use it to derive the confidence intervals for all three distributions in Lemmas 1-3. In subsequent sections we will evaluate the two methods using both Monte Carlo method and real data. In the cases of CSD and CSA we can directly compare the two methods.

3.1 The Fieller Method for CSD and CSA

The CSD (8) is a ratio distribution. The numerator of the CSD is $iS_{i-1}h_i$ and the denominator is $S = \sum_{k=0}^{F} S_k$. The denominator is always above 1 (since $\hat{S}_0$=1). Hence we do not need to worry that the denominator can be close to zero, which might make the CI for a_i inaccurate.

Fieller (1940, 1954) proposed a method to derive the confidence interval of the ratio of two random variables. In Fieller's method, it requires that the numerator and the denominator follow the bivariate normal distribution asymptotically. From observations 1 and 2 above, CSD satisfies this requirement (as does CSA), so we can use Fieller's method to derive the CIs.[5]

We will first derive a test statistic for the CSD estimator, which from (6) can be written in terms of the true values:

$$iS_{i-1}h_i - a_iS = 0$$

If we replace the true values $iS_{i-1}h_i$, and S by their estimators[6] we have the new relationship in terms of the estimators and the true value a_i:

$$i\hat{S}_{i-1}\hat{h}_i - a_i\sum_{k=0}^{F}\hat{S}_k \overset{d.}{\sim} N(0, Var(i\hat{S}_{i-1}\hat{h}_i - a_i\sum_{k=0}^{F}\hat{S}_k)) \tag{12}$$

From (12), we have the test statistic for the null hypothesis for a_i being a certain value (for example, a_i=0).

Theorem 1 : The test statistic for the CSD can be written as:

$$TS_{CSD}(a_i) = \frac{i\hat{S}_{i-1}\hat{h}_i - a_i\hat{S}}{(a_i^2\hat{Var}(\hat{S}) - 2a_i\hat{Cov}(i\hat{S}_{i-1}\hat{h}_i, \hat{S}) + \hat{Var}(i\hat{S}_{i-1}\hat{h}_i))^{\frac{1}{2}}}$$

The TS_{CSD} can be applied to a null hypothesis test with the critical value from the student-t distribution. Clearly, $TS_{CSD} = 0$ when a_i=$\hat{a}_i$ by construction. By applying the same method, we can also derive the test statistic for the CSA estimator.

Corollary 1 : The test statistic for the CSA can be written as:

$$TS_{CSA}(a_i^A) = \frac{\hat{S}_{i-1} - a_i^A\hat{S}}{((a_i^A)^2\hat{Var}(\hat{S}) - 2a_i^A\hat{Cov}(\hat{S}_{i-1}, \hat{S}) + \hat{Var}(\hat{S}_{i-1}))^{\frac{1}{2}}}$$

[5] Franz (2007) provides a guide on deriving the confidence interval of ratio variables using Fieller's method.

[6] In Fieller's theorem, $iS_{i-1}h_i$, and S are replaced by the population mean $E(S_{i-1}h_i)$ and $E(S)$. In reality, we use the sample estimators $\hat{S}_{i-1}\hat{h}_i$ and $\hat{S}$ as the mean value

TS_{CSA} can also be compared with the student-t critical value to perform the null hypothesis test.

Using Theorem 1, we can now derive the CI. If we have a desired confidence level α in terms of the probability of a type 1 error, we have a range of acceptable values of a_i satisfying:

$$P[TS_{CSD}(a_i)] \leq 1 - \alpha/2$$

If $t_{\alpha/2}(d)$ is the critical value in the student-t table with d the degrees of freedom, this implies:

$$i\hat{S}_{i-1}\hat{h}_i - a_i\hat{S} \leq t_{\alpha/2}(d) * (a_i^2 \hat{Var}(\hat{S}) - 2a_i \hat{Cov}(i\hat{S}_{i-1}\hat{h}_i, \hat{S}) + \hat{Var}(i\hat{S}_{i-1}\hat{h}_i))^{\frac{1}{2}} \tag{13}$$

Equation (13) can be written as a quadratic in a_i of the form:

$$F(a_i) = A_{a_i}a_i^2 - 2B_{a_i}a_i + C_{a_i} \leq 0 \tag{14}$$

To derive the confidence interval of a_i, we need to solve out the two roots of equation (14) when $F(a_i) = 0$.

Theorem 2 : The confidence interval of a_i can be defined as the interval between $\lambda_{1,i}$ and $\lambda_{2,i}, which are the roots defined by$:

$$\lambda_{1,i}, \lambda_{2,i} = \frac{B_{a_i} \pm \sqrt{B_{a_i}^2 - A_{a_i}C_{a_i}}}{A_{a_i}} \tag{15}$$

Where:

$$A_{a_i} = \hat{S}^2 - t_{\alpha/2}(d)^2 Var(\hat{S})$$

$$B_{a_i} = i\hat{S}_{i-1}\hat{h}_i\hat{S} - t_{\alpha/2}(d)^2 \hat{Cov}(i\hat{S}_{i-1}\hat{h}_i, \hat{S})$$

$$C_{a_i} = (i\hat{S}_{i-1}\hat{h}_i)^2 - t_{\alpha/2}(d)^2 \hat{Var}(i\hat{S}_{i-1}\hat{h}_i)$$

$\lambda_{1,i}$ is lower bound and $\lambda_{2,i}$ is the upper bound for the estimator of CSD at i-th period. So the true value lies within the interval $\lambda_{1,i} < a_i < \lambda_{2,i}$ with probability $(1 - \alpha)$.

Once again, note that $\hat{S}$, which is the denominator of CSD, is always greater than 1. Hence we do not have to worry about an unbounded CI which would require a different method (see Guiard (1989) and Buonaccorsi (2005) for more details of bounded and unbounded confidence intervals with ratio variables). We can also apply Fieller's method to obtain the CI for CSA.

Corollary 2 : The CI for the CSA estimator can be derived by solving two roots $\lambda_{3,i}$ and $\lambda_{4,i}$ of $F(a_i^A) = A_{a_i^A} a_i^2 - 2B_{a_i^A} a_i + C_{a_i^A} \leq 0$. The CI of a_i^A can be defined as $[\lambda_{3,i}, \lambda_{4,i}]$.

$$\lambda_{3,i}, \lambda_{4,i} = \frac{B_{a_i^A} \pm \sqrt{B_{a_i^A}^2 - A_{a_i^A} C_{a_i^A}}}{A_{a_i^A}} \tag{16}$$

Where:

$$A_{a_i^A} = (\hat{S})^2 - t_{\alpha/2}(d)^2 V\hat{a}r(\hat{S})$$

$$B_{a_i^A} = \hat{S}_{i-1}\hat{S} - t_{\alpha/2}(d)^2 Cov(\hat{S}_{i-1}, \hat{S})$$

$$C_{a_i^A} = (\hat{S}_{i-1})^2 - t_{\alpha/2}(d)^2 Var(\hat{S}_{i-1})$$

Where $\lambda_{3,i}$ is lower bound and $\lambda_{4,i}$ is the upper bound for the estimator of CSA. So we have $\lambda_{3,i} < a_i^A < \lambda_{4,i}$.

3.2 The Delta Method

The delta method (also known as the Taylor expansion method) is an alternative way to find out the CIs for ratio estimators such as CSD and CSA, and can also be applied to non-ratio estimators such as DD. Even if the CSD and CSA estimators do not follow the normal distribution exactly, it is still worth investigating the CIs derived by the delta method since it has proven to be a robust method. We derive the CIs for the CSD, CSA and DD estimators using the standard delta method, which we report in Lemmas 1-3.

Lemma 1 the CI for the CSD estimator derived by the delta method can be written as:

$$\lambda_{1,i}^{del}, \lambda_{2,i}^{del} = \hat{a}_i \pm t_{\alpha/2}(d)\sqrt{\widehat{Var}(\hat{a}_i)} \tag{17}$$

Where $\widehat{Var}(\hat{a}_i)$ is[7]:

$$\widehat{Var}(\hat{a}_i) = i^2 \frac{\widehat{Var}(\hat{S}_{i-1}\hat{h}_i)}{\hat{S}^2} + i^2 \frac{\hat{S}_{i-1}^2 \hat{h}_i^2 \widehat{Var}(\hat{S})}{\hat{S}^4} - 2i^2 \frac{\hat{S}_{i-1}\hat{h}_i \widehat{Cov}(\hat{S}_{i-1}\hat{h}_i, \hat{S})}{\hat{S}^3} \tag{18}$$

Lemma 2 the CI for the CSA estimators derived using the delta method can be written as:

$$\lambda_{3,i}^{del}, \lambda_{4,i}^{del} = \hat{a}_i^A \pm t_{\alpha/2}(d)\sqrt{\widehat{Var}(\hat{a}_i^A)} \tag{19}$$

Where $\widehat{Var}(\hat{a}_1^A)$ is:

$$\widehat{Var}(\hat{a}_i^A) = \frac{\widehat{Var}(\hat{S}_{i-1})}{\hat{S}^2} + \frac{\hat{S}_{i-1}^2 \widehat{Var}(\hat{S})}{\hat{S}^4} - 2\frac{\hat{S}_{i-1}\widehat{Cov}(\hat{S}_{i-1}, \hat{S})}{\hat{S}^3} \tag{20}$$

We can derive the CI for the DD estimators, since they follow the normal distribution asymptotically (see Observation 2).

Lemma 3 the CI for the DD estimator deriving from the delta method can be written as:

$$\lambda_{5,i}^{del}, \lambda_{6,i}^{del} = \hat{a}_i^d \pm t_{\alpha/2}(d)\sqrt{\widehat{Var}(\hat{a}_i^d)} \tag{21}$$

Where $\widehat{Var}(\hat{a}_i^d)$ is:

$$\widehat{Var}(\hat{a}_i^d) = (\hat{S}_{i-1}\hat{h}_i)^2 [\frac{N_i - D_i}{N_i D_i} + \sum_{k=1}^{i-1} \frac{D_k}{N_k(N_k - D_k)}] \tag{22}$$

In the appendix, we also derive the test statistics and CIs for $\hat{S}$ and $\bar{h}$.

4 Monte Carlo Simulations of CIs

In this section we evaluate both the Fieller and delta methods of deriving CIs using a data generating process. In addition, we allow for right-censored observations included in the sample (when some spells are not observed to their end).

[7] the variances of the three estimators of the distributions $\widehat{Var}(\hat{a}_i)$, $\widehat{Var}(\hat{a}_i^A)$ and $\widehat{Var}(\hat{a}_i^d)$ have been derived by Tian and Dixon (2022)

4.1 Data Generating Process

In this section, we describe both the data generating and right-censoring processes. We follow Efron (1981), who used bootstrap method to calculate the variance of KM estimators. He also explained how to simulate survival data (including right-censored observations) and re-sample them. An observation is uncensored when both the start date and the end date are known. An observation is right-censored when the start date is known but the end date is unknown.

The survival time of n-th observation can be expressed as:

$$t_n = min(T_n, C_n) \;\; n = 1, 2, ..., N.$$

where T_n is the life time for n-th observation; C_n is the censored time for n-th observation. We can define the censored coefficient z_n as follows:

$$z_n = \begin{cases} 1 & T_n \leq C_n \\ 0 & C_n < T_n \end{cases}$$

The pairs observations can be defined as:

$$(t_n, z_n) = \begin{cases} (T_n, 1) & min(T_n, C_n) = T_n \\ (C_n, 0) & min(T_n, C_n) = C_n \end{cases}$$

By applying the above rules to construct the new data set, the pairs observations can be defined and included in the new data set as:

$$(t_1, z_1), (t_2, z_2), ..., (t_N, z_N).$$

Before applying the KM estimator to calculate the survival function, the observations need to be re-organized and sorted from the smallest to the largest:$(t_1^s, z_1^s), (t_2^s, z_2^s), ..., (t_N^s, z_N^s)$ where $(t_1^s, z_1^s) = (t_j, z_j)$ depending on $t_j = min(t_1, ..., t_j, ..., t_N)$; $z_1^s = z_j$ corresponds to $t_1^s = t_j$. At this point, the KM estimator can be applied to calculate the survival function at each period $(t_1^s, ..., t_N^s)$. The CPI-micro price-quotes data collected each month and there exist observations in each period (month). To simulate this type of data set, the pairs observations can be located into each time interval (month):$[0, U_1), [U_1, U_2), ..., [U_{F-1}, U_F]$. We use uncensored case as an example. If there are 3 observations $(t_1^s, 1), (t_2^s, 1)$ and $(t_1^s, 1)$ and t_1^s, t_2^s, t_3^s are all greater than 0 and less than U_1, then those three distributions are assigned into $[0, U_1)$ and count $D_1 = 3$ while $N_1 = N - 3$.

We will assume the life-time T and censored-time C of the n-th observation follow the exponential distribution:

$$Pro(T_n > q) = exp(-2q) \tag{23}$$

$$Pro(C_n > p) = exp(-0.5p) \tag{24}$$

Where $p, q \geq 0$. Following (23) and (24), we can simulate the lifetime and censored time and construct the pairs observations:(t_n, z_n). In the uncensored case, where all observations are uncensored, $t_n = T_n$ and $z_n = 1$ for all the observations. If there is no censoring, the survival time of observations T_n is applied to draw the sample directly. With censoring, we need to compare the T_n with C_n. N is the total number of the observations in the sample and F is the maximum length of the period. We only consider right-censored observations, as it is standard for KM estimators to exclude left-censored observations.

Having generated the observations (censored and uncensored), we then assign them into five categories or intervals (months): $[0, 0.1)$, $[0.1, 0.2)$, $[0.2, 0.3)$, $[0.3, 0.5)$ and $[0.5, \infty)$. We use the observations located in interval $[0, 0.1)$ as an example. In uncensored case, the survival function[8] of period $[0, 0.1)$ can be defined as:

$$S_{0.1} = Pro(T_n \geq 0.1) = exp(-0.2)$$

The hazard function is:

$$h_{0.1} = Pro(0 \leq T_n < 0.1 |\ 0 \leq T_n) = 1 - exp(-0.2)$$

Therefore, for the interval $[U_{i-1}, U_i)$ where $i = 1, 2, ..., F$, the true value of the survival function is:

$$S_{U_i} = Pro(T_n \geq U_i) = exp(-2U_i)$$

With $S_{U_0} = 1$ (since $min(T_n, C_n) \geq 0$). The true value of hazard function is:

$$h_{U_i} = \frac{Pro(T_n \geq U_{i-1}) - Pro(T_n \geq U_i)}{Pro(T_n \geq U_{i-1})} = 1 - exp[-2(U_i - U_{i-i})]$$

The true value of DD estimator is:

$$a^d_{U_i} = Pro(U_{i-1} \leq T_n < U_i) = S_{U_{i-1}} h_{U_i}$$

[8]Note that they are not the estimators. They are the true value which is applied to simulate the data set.

In Fieller's method, the CSD estimator at period i is $a_i = \frac{iS_{i-1}h_i}{S}$. So we need to find out the true value of $S_{i-1}h_i$ and S and calculate the value of a_i [9]. The value of the CSD estimator is:

$$a_{U_i} = \frac{i.a^d_{U_i}}{S}$$

Where $S = \sum_{k=U_0}^{U_F} S_k$. The value of the CSA estimator[10] is:

$$a^A_{U_i} = \frac{S_{U_{i-1}}}{S}$$

When the right-censored observations included, the uncensored ratio is $Pro(T_n < C_n) = \frac{2}{2+0.5} = 0.8$ [11] while the censored ratio is 0.2 in the sample. In other words, there are 80% observations are uncensored while 20% of the samples are right-censored. The survival function for the interval $[U_{i-1}, U_i)$ including right-censored observations is:

$$S_{U_i} = Pro(min(T_n, C_n) \geq U_i) = exp(-2.5U_i)$$

The hazard function is:

$$h_{U_i} = Pro(U_{i-1} \leq min(T_n, C_n) < U_i | min(T_n, C_n) \geq U_{i-1}) = [1 - exp(-2.5(U_i - U_{i-1}))]$$

After having the true value of survival function and the hazard function, we can construct the CSD, CSA and DD estimators.

Another common approach which is used is to exclude censored data from the numerator of the hazard function. In this case we have the alternative hazard function $h^*_{U_i}$ where:

$$h^*_{U_i} = Pro(T_n \leq C_n).h_{U_i} = 0.8[1 - exp(-2.5(U_i - U_{i-1}))]$$

We do not use this hazard formulation, not least because when combined with the three distributions, it implies that all censored spells last longer

[9] The mean value (or true value) of ratio variable is very hard to find out when the numerator is correlated with the denominator. In Fieller's theorem, we only need to find out the true value of $S_{i-1}h_i$ and S separately and apply them to construct the benchmark value a_i as the real value of CSD estimator. We also use the similar way to find out the value of a^A_i in simulation.

[10] In CSA estimator, there does not exist $a^A_{U_0}$ due to the definition. It start from $a^A_{U_1}$.

[11] The details of this algebra can be found in probability and stochastic calculus textbooks.

than F periods (or in this case infinity). In effect, by using h_{U_i} we treat censored spells as if they were uncensored when estimating the hazard and survival functions and the three distributions. Censoring simply operates to shorten the spells we observe. However, in Table 5 we provide an illustration using $h^*_{U_i}$.[12]

4.2 The Confidence Interval of Cross-Sectional Distributions

In this section, we investigate the empirical sizes of the Fieller's method and delta method for the CIs of the estimators of CSD, CSA and DD. The purpose is that we want to find out whether both Fieller's method and delta method are valid for the CI of CSD and CSA estimators. We also test whether the CI calculated from delta method of DD estimator is valid (accuracy of delta method for the CI).

To evaluate the empirical size, the idea is that we evaluate whether the true values are located in their CI in each simulation for each interval $[U_{i-1}, U_i), i = 1, 2, ..., F$. After 100,000 simulations, we count how many times the true values ($\frac{iS_{i-1}h_i}{S}$, $\frac{S_i}{\sum S}$ and $S_{i-1}h_i$) are included in their CI which gives the empirical size. If the empirical size is close to α, then it means the CI is accurate.

Step 1: We use the data generated according to section 4.1 to calculate CSD, CSA and DD estimators over the five duration categories, $i = 1...F$. The sample sizes for each simulation are either $N = 50$, $N = 100$, $N = 200$ and $N = 400$. Since we do not know the exact distribution of the CSD, CSA and DD[13], to evaluating the empirical sizes of those methods, we investigate whether the "true" values of CSD and CSA are located in their CI calculated using Fieller's method and the delta method; and for the DD estimator if $S_{i-1}h_i$ is located in the CI calculated from the delta method. The significant levels chosen are the standard ones α (type I error) being $10\%, 5\%$ and 1%.

[12]For example, in medical applications there might be interest in the occurrence of a very specific outcome (heart attack, death) rather than monitoring what happens in cases where this does not occur.

[13]The DD estimator should follow the normal distribution asymptotically. At this point, we want to use the DGP to double check whether the DD estimator follows the normal distribution asymptotically. In other words, does the delta method provides accurate CIs for DD.

Step 2: Repeat the step 1 by $M = 100,000$ times. This generates $100,000$ CIs for each of the CSD, CSA and DD estimators for each of the 5 categories (time periods) in the different sample sizes. We can evaluate the number of times when the true values are included in the estimated CIs. Furthermore, if the empirical size is close to α, it means that the CI is accurate.

4.2.1 The Results without Censoring

Table (1), (2) and (3) show the empirical size of the three estimators without censoring. Table (1) lists the empirical sizes for the CSD. In the first column, it shows the categories of the CSD estimators. In terms of the DGP, the observations are assigned into five categories: $[0, 0.1)$, $[0.1, 0.2)$, $[0.2, 0.3)$, $[0.3, 0.5)$ and $[0.5, \infty)$. For example, $a_{0.1}$ means the CSD estimator of category $[0, 0.1)$; a_{inf} means the CSD estimator of category $[0.5, \infty)$. In other words, $a_{0.1}$ can be also known as CSD estimator for the first period; a_{inf} can be also known as CSD estimator for the fifth period. When we set the theoretically empirical size is 10 % with N=50, it can be seen that the empirical sample size of CI from Fieller's method is superior to delta method except in the first period. When the theoretically empirical size is 5 %, the 5 periods CIs from Fieller's method are closer to 5 % compared with delta method. With respect to 1 %, we still have the same result that CI from Fieller's method are closer to 1 % for all periods. When the sample size increased to 100 and 200, the performances of Fieller's method and delta method are nearly the same when α=10 % and 5 %. In the case of 1 %, it can be seen that the empirical size calculated from Fieller's method are closer to 1 % for all periods compared with delta method. When the sample size increased to 400 and α=10 %, it can be seen that empirical size of Fieller's method is superior to delta method for $a_{0.1}$, $a_{0.2}$, $a_{0.3}$ and a_{inf}. With respect to N=400 and α= 5 %, the empirical size of Fieller's method is closer to 5 % compared with delta method. In terms of 1 %, empirical size of all the 5 periods CI of CSD from Fieller's method is superior to delta method. Therefore, we can make the a conclusion that the CI of CSD calculated by Fieller's method provides a more accurate empirical size when all the observations are uncensored. In addition, even though the delta method is inferior to Fieller's method, it is still a valid method to calculate the CI for CSD.

Table (2) provides the empirical size for the CSA estimators. When N=50, both Fieller's method and delta method give a similar empirical size for CSA estimators even though empirical results are slightly biased, particularly for the 1% case, where both methods reject too many true values. For larger sample sizes (N=100, 200 and 400), the empirical sample size is quite accurate for all significance levels 10%, 5% and 1%. In addition, there are no significant differences of empirical size calculated from Fieller's method and the delta method. We also calculate the empirical size of $\hat{S}$, for which we only use the delta method. Again, this is quite accurate in larger sample sizes, and in small sample $N = 50$ performs the worst for the 1% significance level.

Table (3) describes the empirical size for the DD estimators using the delta method. When the sample size is 50, it can be seen that the empirical size is acceptable when the type I error is 10% except $a^d_{0.2}$ and $a^d_{0.3}$. In terms of the 5%, the empirical result is quite good even the sample size is small. While the type I error is chosen to be 1%, the empirical size is slightly far from 1% except the last period of DD estimator a^d_{inf}. When the sample size increases to 100, 200 and 400, all the empirical size of the DD estimators tend to be very close to the theoretically empirical size (10%, 5% and 1%). In other words, we use the DGP to show that the delta method is valid for DD estimator. The empirical size is improved with the increased sample size.

4.2.2 The Results with Censoring

Table (4), (5) and (6) show the empirical size of the three estimators when there is censoring of spells as outlined previously.

Table (4) shows the empirical size of the CSD when the censored observations are included in the sample. As we can seen when the α=10 % and N=50, the empirical size of $a_{0.2}$, $a_{0.3}$, $a_{0.5}$ and a_{inf} calculated from Fieller's method are slightly closer to 10 % while $a_{0.1}$ calculated from delta method is closer to 10 %. When α=5 %, all the five periods of CSD from Fieller's method are slightly superior to delta method. When α=1 %, the empirical size is significantly improved and closer to 1 % calculated from Fieller's method. When the sample size increase to 100 and 200, Fieller's method is still better than delta method even both methods provide quite accurate empirical size. When the sample size increase to 400 with α=10 % and 5 %, both methods provide quite similar empirical size. With respect to α=1 %, Fieller's method is superior to delta method. In conclusion, empirical size of the CSD calculated from Fieller's method is more accurate when there are censored observations. However, delta method is also valid since it still provides an acceptable empirical size.

With respect to CSA estimators including the censored observations, the empirical result can be found in table (5). Both methods give a quite accurate empirical size even the sample size N=50 while the empirical size of the CSA estimator in last period is a little far from the theoretically empirical size. With the increase of the sample size, the empirical size of the CSA estimator of the last period is significantly improved. Especially, the empirical size is quite accurate for all the CSA estimators when N=400. Note that we have an additional coefficient included, $a^{A*}_{\inf+1}$, which is calculated by using the alternative and commonly used hazard formulation $h^*_{U_i}$. We provide this merely as an illustrative example and we do not use it elsewhere in the paper.[14]

Table (6) gives the empirical size for DD estimator calculated by the delta method. For the small sample (N=50), the empirical size of DD is far from the theoretically empirical size (α=10%, 5% and 1%). However, with larger samples, the empirical size of DD estimators in all 5 periods are close to the theoretically empirical size. In other words, the delta method is valid for CI of DD estimators depending on the simulation.

5 Application to CPI Micro-Price Data

From the Monte Carlo simulations, we can see that whilst both Fieller's method and the delta method can be applied to calculate the CI of the CSD, Fieller's method can potentially improve the empirical size. In this section, we apply both methods to calculate CI of the CSD and CSA and the delta method for DD using the UK CPI micro-price data.[15] Since this is a big data set, given the simulation results we would expect our CIs to be accurate and similar for both methods.

[14] In some applications, a particular endpoint is the focus. For example, in medical studies, the occurrence of a heart attack might be the end point. This is captured by $h^*_{U_i}$, and there is no interest in spells where no heart attack is observed.

[15] In an earlier version of the paper we also used data on UK health service waiting time data, as used in Dixon and Siciliani (2009)]

5.1 Description of the UK CPI Price-Quote Data

We use the UK micro-price data used to construct the monthly CPI inflation statistics. The CPI micro data is obtained from the UK Office for National Statistics (ONS). This data set gives over 100,000 price-quotes each month across over 700 items sampled from different sellers across the UK in order to measure CPI inflation: it includes over 30 million price-quotes from December 1998 to January 2017. These monthly price quotes are collected "locally" by the ONS officers (other prices which are collected centrally are not included in the data set).

From the price-quotes, we can construct price-spells: these are the sequences of monthly quotes where an individual price-setter sets the same price for a particular item each month. These spells are sorted into durations of 1-61 months. For $i = 1, 2, ..., 61$, a price spell has duration i if it lasts exactly i months. For example if the duration is 12 months, that means we observe twelve consecutive months where the seller set the same price and in both the preceding and following month set a different price. The only exception is 61 months, where all spells of 61 months or longer are counted. It is common practice to truncate the distribution in this way. For applications of this type of data and distributions in dynamic macroeconomic models, see for example Dixon and Bihan (2012) using French CPI data and Dixon (2012) using the UK data. Following these two papers, we do not use left-censored spells and assume that right-censored spells end with price-changes. Furthermore, we exclude sales from the data. There are some standard issues with the CPI data, and here we simply follow the normal conventions (see for example Cavallo (2016), Nakamura and Steinsson (2008), Dixon and Tian (2017) and Dixon et al. (2020)). A description of the data and more detailed discussion of methods is included in the Appendix 7.2.

5.2 Empirical Results

First we use Theorem 1 to perform a null hypothesis test on CSD which we report for a range of durations in table (7). The null hypothesis is that the coefficients are equal to the mean coefficient across all durations i: $\hat{a}_i = E(\hat{a}_i) = 0.0164$. In other words, we want to test whether the CSD estimators are equal to their collective arithmetic mean in each period i (if it were accepted for all durations, then we wold have a uniform distribution). As can be seen from table (7), all test statistics of all the CSD estimators are significant and reject the null hypothesis.[16]. In addition, we also report the null hypothesis tests for $\hat{S} = 10.75$ and $\bar{h} = 0.0955$. The value tested for $\bar{h}$ is the arithmetic mean hazard across all durations, and the value for $\hat{S}$ is its reciprocal. Both test statistics clearly reject the null as we would expect. Having shown how to use Theorem 1 (Corollary 1 for $\bar{h}$, Lemma 3 for $\hat{S}$) to preform a null hypothesis test, we now move on to use we Theorem 2 and Lemma 1 to construct the CIs for CSD estimators, along with CSA, DD, $\bar{h}$ and S.

In Figure 1 we show the estimated coefficients for the three distributions in months 1-60 (the full CI results for months 1-61 are given in the online appendix for each distribution). We do not depict the CIs, as these are very small due to the large sample size. However, we can compare the *precision* of the estimates, being the ratio of the absolute size of the CI relative to the coefficient, with a higher value representing a less precise estimate. These are depicted in Figure 2 for months 1-60. We can see that the precision of CSD and DD estimates are almost the same after a few months. In month 1 the precision of CSD is 0.44% whilst DD is 0.28%. However, in later months the difference in precision is very small (less than 0.2% in absolute terms). In general, as time goes on the CI increases relative to the estimate: from below 0.5% in the first months, to over 10% for CSD and DD. The fact that the DD and CSD have such similar levels of precision reflects the fact that the CSD estimator is equal to the product of constant i and $\bar{h}$ times the corresponding DD estimator. The estimator $\bar{h}$ has a very small variance realtive to the DD estimator. Hence precision of DD and CSD are very similar. The fall in precision as i increases reflects the fact that the number of spells and price changes (D and N) become smaller over time. CSA is significantly more precise than CSD and DD, although it too becomes less precise for longer durations. The reason for the higher precision of CSA is because the variance of $\hat{S}_i$ is less than the variance of $\widehat{S_{i-1}}\hat{h}_i$, since $\hat{h}_i$ is highly variable. In the online appendix, we show the CIs of the three distributions for months 1-60.

Table(8) shows a selection of coefficients across each of the 5-years CSD and its CIs calculated from Fieller's method to 4 decimal places, which are the same for the delta method.[17] The reason is the large large number of

[16]Although we do not report the values, this null hypothesis is rejected for all the estimated coefficients $i = 1...61$

[17]The differences exist after 7 decimal places.

observations, which makes the delta method robust for CSD estimators. As we can see, the selected coefficients whilst not monotonic are tending to become smaller (from 0.057 when i=2 to 0.002 for i=52). There is a long fat tail, with 8% of price spells lasting more than 48 months and 6% longer than 60 months. All the CIs in table (8) exclude 0, and so all the CSD estimators are significantly different from 0. The Table also reports the first 6 months of CSA and DD, and at the bottom the estimate and CI of $\hat{S}$, with the results for $\bar{h}$ given by $\hat{a}_1^A$. Note that this estimate of $\bar{h}$ is similar to that reported in Dixon and Tian (2017) which used a slightly different time period.

Whilst the focus of this paper is on evaluating the estimates of the CIs of the three distributions, it is worth noting that the estimates in the online appendix themselves satisfy the identities we earlier outlined earlier. In particular, we can note that from equations (3) and (7) we have:

$$\hat{a}_i^A = \bar{h} = 0.1733 \tag{25}$$

Our estimate of $\bar{S}$ above implies that the average duration of a price spell is just under 6 months, and that the two distributions CSD and DD "cross" between 5 and 6 months, which we see from Table 8 is indeed the case:

$$\hat{a}_5^d = 0.056 > 0.049 = \hat{a}_5$$

$$\hat{a}_6^d = 0.046 < 0.048 = \hat{a}_6$$

These two examples illustrate how the estimators we are using satisfy the theoretical relationships that hold for the underlying distributions (see appendix 7.2 for some further key relationships). It remains a further topic for research to see if we can link the CIs for the different distributions in a similar manner and possibly improve the CIs by deriving them jointly.

6 Conclusion

CSD is a new estimator which has been applied in economics and survival analysis. By applying the theoretical results of Tian and Dixon (2022), we provide the analytic formulae of the confidence interval for the CSD, CSA and DD estimators. The CIs of the CSD and CSA estimator can be calculate from Fieller's method and delta method while the CI of DD estimator can be derived from delta method. In this paper, the Filler's method and the delta method are applied to derive the CIs of the CSD and CSA estimators (delta method for DD estimator) and to derive associated test statistics. To compare the CIs from the two methods, we use the Monte Carlo simulations to evaluate the empirical size. We simulate 5 period CSD, CSA and DD estimators with the sample size N=50, 100, 200 and 400 with type I error = 10 %, 5 % and 1 %. The empirical results of the simulations indicate that Fieller's method is superior to the delta method whether censored observations are included or excluded. However, the delta method still provide acceptable CIs for CSD and CSA estimators in large samples. Furthermore, the CIs of DD and S and $\bar{h}$ estimators derived from the delta method are also accurate.

Finally, we use both methods in real data sets: the UK CPI micro price quotes data. We provide estimates of CSD estimators,the CIs and test statistics, and similarly for CSA, DD, S and $\bar{h}$. In all cases we have very precise estimates of the coefficients, with small CIs which exclude zero. Even very long durations of up to 60 months have significant coefficients despite the small proportion of spells surviving that long. Due to the large sample size, both the Fieller and delta methods yield almost the same CIs, with only minor differences.

We believe that these results provide a solid foundation for using the CSD estimator, and that it will go on to have many more applications in different contexts. Furthermore, the linking together of three distributions allows us to estimate all three distributions from different types of data: if you can estimate one, you can estimate them all.

7 Appendix

7.1 Derivation of CI and test statistics for $\hat{S}$ and $\bar{h}$

First recall that we have the expression for the variance of the sum of survival probabilities:

$$Var(\hat{S}) = Var(\hat{S}_0 + \hat{S}_1 + ... + \hat{S}_F) = \sum_{k=1}^{F} Var(\hat{S}_k) + 2\sum_{k\neq j}^{F} Cov(\hat{S}_k, \hat{S}_j)$$

By applying the Greenwood formula and the covariance formula as in equations (9) and (10), we can calculate the variance of $\hat{S}$. Turning to $Var(\bar{h})$, we could develop an expression using the definition:

$$Var(\bar{h}) = Var(1/\hat{S})$$

However, it is not easy to use this definition to derive $Var(\bar{h})$. Instead, we can use the delta method from Lemma 2 and substitute $\alpha_1^A = \bar{h}$ into equation (20). Noting that since $\hat{S}_0 = 1$, we have $\hat{Cov}(\hat{S}_0, \hat{S}) = 0$ and $\hat{Var}(\hat{S}_0) = 0$. Hence (20) simplifies to:

$$Var(\bar{h}) = \frac{Var(\hat{S})}{\hat{S}^4} \tag{26}$$

We can now calculate the CI for $\hat{S}$ and $\bar{h}$. Since the summation $\hat{S}$ follows the multi-normal distribution asymptotically (from observations 1 and 2), we can write its CI in the standard form:

$$\hat{S} - t_{1-\alpha/2}(d)\sqrt{\hat{Var}(\hat{S})} < S < \hat{S} + t_{1-\alpha/2}(d)\sqrt{\hat{Var}(\hat{S})}$$

The corresponding hypothesis test statistic is therefore:

$$TS_S = \frac{S - \hat{S}}{\sqrt{\hat{Var}\hat{S}}}$$

If we set the null hypothesis $H_0 : \hat{S} = X$, we then have the test for whether $\hat{S}$ is significantly different from X.

$$TS_S = \frac{\hat{S} - X}{\sqrt{\hat{Var}\hat{S}}} \tag{27}$$

Recall that since $S_0 = 1$, the estimator $\hat{S}$ cannot be less than 1. When S=1, we have the case where no spells last more than 1 period. In the case of the price data this would mean that all prices change each period. The test of null hypothesis $H_0 : S = 1$ would then be a test that there was no nominal rigidity in the economy.

Turning to $\bar{h}$, by analogous arguments we have the test statistic:

$$TS_{\bar{h}} = \frac{\bar{h}^* - \bar{h}}{\hat{Var}(\bar{h})^{\frac{1}{2}}} \tag{28}$$

where the variance is given by (26). The CI takes the form:

$$\bar{h} - t_{1-\alpha/2}(d)\sqrt{\hat{Var}(\bar{h})} < \bar{h}^* < \bar{h} + t_{1-\alpha/2}(d)\sqrt{\hat{Var}(\bar{h})}$$

7.2 Relationships between Estimators

There are many direct links between the estimators of the three distributions DD, CSD and CSA and the Survival function S_i and hazard function h_i, which are described in detail in Tian and Dixon (2022) (Table 1). These relationships include:

$$\bar{h} = \sum_{i=1}^{F} \frac{\hat{a}_i}{i} \tag{29}$$

$$\bar{h} = \frac{1}{\sum_i^F i.\hat{a}_k^d} \tag{30}$$

$$\hat{S}_i = 1 - \sum_{k<(i+1)} \hat{a}_k^d \tag{31}$$

$$\hat{a}_i = i(\hat{a}_i^A - \hat{a}_{i+1}^A) \tag{32}$$

We can also derive a new result using (31),

$$\hat{S} = 1 + \sum_{k=2}^{F} \hat{a}_k^d + \sum_{k=3}^{F} \hat{a}_k^d .... = \sum_{k=1}^{F} (F - k + 1)\hat{a}_k^d \tag{33}$$

7.3 Description of CPI-Data

We use the CPI micro data obtained from the UK Office for National Statistics (ONS), the UK government's statistics department. The data includes over 30 million price quotes from December 1998 to January 2017. These monthly price quotes are collected "locally" by the ONS officers. The duration of the price spell means that how long the price set by a specific retailer in a specific region for a specific item spell lasts, in terms of a sequence of identical price-quotes across consecutive months.

In terms of the UK CPI micro data, the observations are divided into 13 divisions which are associated with the Classification of Individual Consumption According to Purpose (COICOP). The COICOP codes including 5-6 digit numbers are applied to classify the goods and the services. After 2015, ONS also reports the COICOP5 codes which give more details for observations. The locally collected data which we were able to use covers 11 of the two digit COICOP divisions. There are multi-observations for the same item, since the prices are collected from different shops and regions to make a representative sample. There are 12 regions used to construct the data, including London, Scotland, Wales and Northern Ireland etc. There also exist the shop codes for the different brands of shops such as Tesco and Sainsbury.

There exist the weighted coefficients for the shop, the region, the item and the COICOP division (COICOP weight). The weighted coefficients can be applied to calculate the weighted price (aggregate price). Therefore, the different levels of the indices can be generated. The weighted coefficients of different regions are named as stratum weight in the data.

There exist some indicators for the price quotes. An indicator "S" gives the information that the observation is on "sale". In addition, if the price quote of observation is missing, an indicator factor "M" is assigned to the observation. An indicator "T" is assigned to the observation when the observation is temporary out of the stock.

There exist the records of the start date and end date for each observation. This is a very useful information when the censored and truncation problems are considered. Currently, some observations are still tracked and included in the CPI framework so that the end date is 999999. Since ONS may change the CPI framework each year, some goods and services may not be included in the CPI framework. In such cases, old items are replaced by new items.

For each observation, both the price and log price are given in the data. Those values can be applied to calculate the frequency and the size of the price changes. There are sometimes gaps, where the price quotes are missing for some period and reappear later. The ONS uses the missing indicator is "M" for these cases. There are several methods to deal with this problem. If the price quote is missing in a special period, the price can be assumed as the same as the its last record until the price quote appears again. See Cavallo (2016) for an evaluation of this approach and comparison with alternatives considered in Nakamura and Steinsson (2008). Here we use the last record for the missing price quotes.

In addition, some price-quotes are sales, temporary or terminal discounts. Many studies argue that the data including sales may affect the economic relevance of the duration and frequency, so we follow the common practice and replace the sale price by its previous record (the same method used to deal with missing price). Another issue is "outliers": some price quotes may be increased by 200 percent or 300 percent in one month. In reality, this is more likely to be a mis-measurement rather than an actual like-for-like price rise. If prices increased more than 130 percent month on month, they are deleted from the data. Likewise, if the prices are decreased more than 75 percent, they are also deleted from the data. (See Nakamura and Steinsson (2008), Dixon and Tian (2017) and Dixon et al. (2020) for a discussion of this and other data issues)

Table 1: The Empirical Size of CSD

All the Observations Are Uncensored.

CSD	Fieller's Method			Delta Method		
N=50	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.11706	0.06797	0.02880	0.10930	0.07703	0.03588
$a_{0.2}$	0.12486	0.06902	0.03148	0.12463	0.07383	0.03877
$a_{0.3}$	0.13422	0.07447	0.04167	0.13588	0.07994	0.04454
$a_{0.5}$	0.11700	0.06844	0.02502	0.11945	0.07104	0.02731
a_{inf}	0.11139	0.05849	0.01520	0.11061	0.06077	0.01607
N=100	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.10798	0.05931	0.01822	0.10723	0.06146	0.02115
$a_{0.2}$	0.11185	0.06153	0.01980	0.11110	0.06327	0.02267
$a_{0.3}$	0.11475	0.06641	0.02288	0.11575	0.06843	0.02520
$a_{0.5}$	0.10902	0.05843	0.01626	0.11009	0.05990	0.01716
a_{inf}	0.10379	0.05459	0.01235	0.10448	0.05536	0.01325
N=200	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.10436	0.05461	0.01386	0.10490	0.05543	0.01629
$a_{0.2}$	0.10579	0.05607	0.01447	0.10651	0.05644	0.01593
$a_{0.3}$	0.10611	0.05771	0.01590	0.10645	0.05847	0.01683
$a_{0.5}$	0.10594	0.05547	0.01331	0.10632	0.05618	0.01363
a_{inf}	0.10151	0.05096	0.01079	0.10208	0.05179	0.01098
N=400	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.10238	0.05279	0.01117	0.10236	0.05272	0.01235
$a_{0.2}$	0.10204	0.05323	0.01285	0.10242	0.05371	0.01354
$a_{0.3}$	0.10506	0.05608	0.01327	0.10505	0.05690	0.01389
$a_{0.5}$	0.10169	0.05125	0.01103	0.10189	0.05162	0.01119
a_{inf}	0.10083	0.05111	0.01045	0.10107	0.05114	0.01048

Table 2: The Empirical Size of CSA and S
All the Observations Are Uncensored.

CSA	Fieller's Method			Delta Method		
N=50	10%	5%	1%	10%	5%	1%
$a_{0.1}^A$	0.10692	0.05872	0.01881	0.10734	0.05934	0.01876
$a_{0.2}^A$	0.11187	0.06070	0.01730	0.10705	0.05756	0.01631
$a_{0.3}^A$	0.10803	0.05657	0.01364	0.10161	0.05087	0.01059
$a_{0.5}^A$	0.10912	0.05862	0.01437	0.10770	0.05907	0.01738
a_{inf}^A	0.11139	0.05849	0.01520	0.11061	0.06077	0.01607
S	-	-	-	0.10953	0.05852	0.01528
N=100	10%	5%	1%	10%	5%	1%
$a_{0.1}^A$	0.10298	0.05407	0.01437	0.10327	0.05390	0.01438
$a_{0.2}^A$	0.10539	0.05646	0.01429	0.10331	0.05424	0.01381
$a_{0.3}^A$	0.10564	0.05460	0.01297	0.10184	0.05163	0.01152
$a_{0.5}^A$	0.10401	0.05419	0.01237	0.10339	0.05459	0.01367
a_{inf}^A	0.10379	0.05459	0.01235	0.10448	0.05536	0.01325
S	-	-	-	0.10416	0.05374	0.01248
N=200	10%	5%	1%	10%	5%	1%
$a_{0.1}^A$	0.10207	0.05305	0.01223	0.10176	0.05307	0.01230
$a_{0.2}^A$	0.10114	0.05220	0.01226	0.10012	0.05155	0.01218
$a_{0.3}^A$	0.10135	0.05204	0.01103	0.10007	0.05042	0.01039
$a_{0.5}^A$	0.10347	0.05334	0.01128	0.10320	0.05271	0.01210
a_{inf}^A	0.10151	0.05096	0.01079	0.10208	0.05179	0.01098
S	-	-	-	0.10297	0.05253	0.01112
N=400	10%	5%	1%	10%	5%	1%
$a_{0.1}^A$	0.09976	0.05094	0.01088	0.09966	0.05088	0.01088
$a_{0.2}^A$	0.10162	0.05112	0.01161	0.10111	0.05075	0.01138
$a_{0.3}^A$	0.10178	0.05140	0.01045	0.10083	0.05093	0.01024
$a_{0.5}^A$	0.10257	0.05189	0.01023	0.10225	0.05189	0.01080
a_{inf}^A	0.10083	0.05111	0.01045	0.10107	0.05114	0.01048
S	-	-	-	0.10066	0.05074	0.01034

Table 3: The Empirical Size of DD
All the Observations Are Uncensored.

DD	Delta Method		
N=50	10%	5%	1%
$a^d_{0.1}$	0.11476	0.05060	0.04012
$a^d_{0.2}$	0.14394	0.05930	0.03939
$a^d_{0.3}$	0.16189	0.06179	0.04783
$a^d_{0.5}$	0.11701	0.05094	0.04060
a^d_{inf}	0.11067	0.05840	0.01232
N=100	10%	5%	1%
$a^d_{0.1}$	0.09859	0.05451	0.02117
$a^d_{0.2}$	0.09494	0.07047	0.01469
$a^d_{0.3}$	0.10292	0.08796	0.03554
$a^d_{0.5}$	0.09875	0.05476	0.02085
a^d_{inf}	0.11943	0.04810	0.01275
N=200	10%	5%	1%
$a^d_{0.1}$	0.10677	0.04689	0.01470
$a^d_{0.2}$	0.11640	0.06549	0.01215
$a^d_{0.3}$	0.10624	0.05727	0.01384
$a^d_{0.5}$	0.10982	0.04871	0.01473
a^d_{inf}	0.09074	0.04690	0.01255
N=400	10%	5%	1%
$a^d_{0.1}$	0.09411	0.05327	0.01019
$a^d_{0.2}$	0.09455	0.05117	0.01422
$a^d_{0.3}$	0.09675	0.06276	0.01589
$a^d_{0.5}$	0.10429	0.05260	0.01065
a^d_{inf}	0.09643	0.04874	0.01089

Table 4: The Empirical Size of CSD Estimators

Censored Observations are included.

CSD	Fieller's Method			Delta Method		
N=50	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.12561	0.06816	0.02796	0.12029	0.07627	0.03920
$a_{0.2}$	0.11852	0.07538	0.02887	0.12049	0.07646	0.03586
$a_{0.3}$	0.12470	0.08555	0.03445	0.12633	0.08643	0.04173
$a_{0.5}$	0.12113	0.07295	0.02976	0.12365	0.07571	0.03280
a_{inf}	0.10836	0.05978	0.01652	0.11110	0.06201	0.01787
N=100	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.11175	0.06162	0.01865	0.10949	0.06318	0.02245
$a_{0.2}$	0.11248	0.06281	0.02068	0.11185	0.06465	0.02365
$a_{0.3}$	0.11582	0.06767	0.02410	0.11627	0.06919	0.02679
$a_{0.5}$	0.10965	0.06042	0.01822	0.11077	0.06209	0.01950
a_{inf}	0.10645	0.05527	0.01263	0.10877	0.05741	0.01375
N=200	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.10787	0.05733	0.01530	0.010726	0.05772	0.01714
$a_{0.2}$	0.10558	0.05590	0.01498	0.10539	0.05744	0.01669
$a_{0.3}$	0.10869	0.05864	0.01669	0.10885	0.05920	0.01788
$a_{0.5}$	0.10573	0.05554	0.01440	0.10650	0.05625	0.01496
a_{inf}	0.10485	0.05443	0.01228	0.10616	0.05513	0.01293
N=400	10%	5%	1%	10%	5%	1%
$a_{0.1}$	0.10514	0.05431	0.01226	0.10489	0.05481	0.01333
$a_{0.2}$	0.10207	0.05257	0.01245	0.10202	0.05325	0.01343
$a_{0.3}$	0.10465	0.05420	0.01339	0.10494	0.05488	0.01419
$a_{0.5}$	0.10383	0.05318	0.01176	0.10397	0.05354	0.01204
a_{inf}	0.10165	0.05084	0.01105	0.10216	0.05139	0.01134

Table 5: The Empirical Size of CSA and S

Censored Observations are included.

CSA	Fieller's Method			Delta Method		
N=50	10%	5%	1%	10%	5%	1%
$a^A_{0.1}$	0.10215	0.05637	0.01786	0.10231	0.05685	0.01842
$a^A_{0.2}$	0.10397	0.05697	0.01464	0.10313	0.05606	0.01591
$a^A_{0.3}$	0.10618	0.05463	0.01319	0.09893	0.04853	0.00970
$a^A_{0.4}$	0.10823	0.05839	0.01481	0.10537	0.05707	0.01617
$a^A_{\inf}$	0.10851	0.05944	0.01535	0.11137	0.06163	0.01773
$a^A_{\inf+1}*$	0.13182	0.06381	0.00263	0.13299	0.06275	0.00271
S	-	-	-	0.10293	0.05371	0.01367
N=100	10%	5%	1%	10%	5%	1%
$a^A_{0.1}$	0.10255	0.05408	0.01481	0.10247	0.05416	0.01496
$a^A_{0.2}$	0.10426	0.05501	0.01289	0.10280	0.05420	0.01306
$a^A_{0.3}$	0.10521	0.05349	0.01235	0.10204	0.05044	0.01039
$a^A_{0.5}$	0.10478	0.05408	0.01289	0.10399	0.05435	0.01365
$a^A_{\inf}$	0.10407	0.05428	0.01276	0.10440	0.05522	0.01351
$a^A_{\inf+1}*$	0.12890	0.08086	0.04062	0.12987	0.08133	0.03894
S	-	-	-	0.10364	0.05455	0.01256
N=200	10%	5%	1%	10%	5%	1%
$a^A_{0.1}$	0.10164	0.05164	0.01203	0.10168	0.05173	0.01208
$a^A_{0.2}$	0.10325	0.05304	0.01154	0.10222	0.05289	0.01189
$a^A_{0.3}$	0.10222	0.05173	0.01127	0.10030	0.05001	0.01018
$a^A_{0.5}$	0.10278	0.05187	0.01134	0.10209	0.05188	0.01212
$a^A_{\inf}$	0.10338	0.05350	0.01174	0.10371	0.05354	0.01212
$a^A_{\inf+1}*$	0.11641	0.06840	0.02515	0.11708	0.06833	0.02427
S	-	-	-	0.10272	0.05191	0.01084
N=400	10%	5%	1%	10%	5%	1%
$a^A_{0.1}$	0.10179	0.05099	0.01123	0.10177	0.05103	0.01126
$a^A_{0.2}$	0.10155	0.05148	0.01141	0.10098	0.05121	0.01144
$a^A_{0.3}$	0.10048	0.05036	0.00995	0.09943	0.04943	0.00964
$a^A_{0.5}$	0.10197	0.05155	0.01109	0.10184	0.05201	0.01144
$a^A_{\inf}$	0.10144	0.05186	0.01141	0.10122	0.05191	0.01141
$a^A_{\inf+1}*$	0.10831	0.05844	0.01729	0.10885	0.05824	0.01679
S	-	-	-	0.10214	0.05094	0.01085

Table 6: The Empirical Size of DD Estimator
Censored Observations are included.

DD	Delta Method		
N=50	10%	5%	1%
$a^d_{0.1}$	0.12502	0.05482	0.01737
$a^d_{0.2}$	0.10530	0.07979	0.02418
$a^d_{0.3}$	0.11625	0.08935	0.03371
$a^d_{0.5}$	0.12385	0.07481	0.03331
a^d_{inf}	0.11304	0.06395	0.01955
N=100	10%	5%	1%
$a^d_{0.1}$	0.12259	0.05928	0.01290
$a^d_{0.2}$	0.11444	0.06556	0.02251
$a^d_{0.3}$	0.11480	0.07002	0.02702
$a^d_{0.5}$	0.11080	0.06112	0.01970
a^d_{inf}	0.10659	0.05619	0.01353
N=200	10%	5%	1%
$a^d_{0.1}$	0.10273	0.06141	0.01389
$a^d_{0.2}$	0.10578	0.05687	0.01573
$a^d_{0.3}$	0.10908	0.05960	0.01739
$a^d_{0.5}$	0.10628	0.05649	0.01507
a^d_{inf}	0.10540	0.05449	0.01199
N=400	10%	5%	1%
$a^d_{0.1}$	0.10607	0.06041	0.01275
$a^d_{0.2}$	0.10197	0.05235	0.01266
$a^d_{0.3}$	0.10489	0.05474	0.01370
$a^d_{0.5}$	0.10319	0.05358	0.01229
a^d_{inf}	0.10161	0.05200	0.01116

Table 7: The Test Statistics Calculated from Theorem 2 for CSD Estimators of UK Micro-CPI Data from 1998m12 to 2017m1. The Null Hypothesis is: $\hat{a}_i = E(\hat{a}_i) = 0.01639344$

$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$
-629.96	-569.72	-561.64	-550.47
$\hat{a}_5$	$\hat{a}_6$	$\hat{a}_7$	$\hat{a}_8$
-587.13	-579.87	-598.88	-587.01
$\hat{a}_9$	$\hat{a}_{10}$	$\hat{a}_{11}$	$\hat{a}_{12}$
-624.34	-631.83	-603.61	-544.72
$\hat{a}_{13}$	$\hat{a}_{14}$	$\hat{a}_{15}$	$\hat{a}_{16}$
-674.64	-705.80	-720.86	-711.95
$\hat{a}_{25}$	$\hat{a}_{26}$	$\hat{a}_{27}$	$\hat{a}_{28}$
-788.83	-805.47	-806.84	-788.51
$\hat{a}_{37}$	$\hat{a}_{38}$	$\hat{a}_{39}$	$\hat{a}_{40}$
-857.54	-854.13	-863.24	-854.20
$\hat{a}_{49}$	$\hat{a}_{50}$	$\hat{a}_{51}$	$\hat{a}_{52}$
-879.51	-888.84	-892.93	-881.19
$\hat{S}$	$\bar{h}$		
-8.76	8.68		

Table 8: The 90% Confidence Interval of Different Estimators for CPI Data From 1998m12 to 2017m1

CSD	Fieller's Method	Precision
$\hat{a}_1$=0.0491	[0.0490, 0.0493]	0.50%
$\hat{a}_2$=0.0572	[0.0570, 0.0574]	0.56%
$\hat{a}_3$=0.0565	[0.0563, 0.0566]	0.64%
$\hat{a}_4$=0.0564	[0.0562, 0.0566]	0.70%
$\hat{a}_5$=0.0486	[0.0484, 0.0488]	0.81%
$\hat{a}_6$=0.0481	[0.0479, 0.0483]	0.87%
$\hat{a}_7$=0.0437	[0.0435, 0.0439]	0.97%
$\hat{a}_8$=0.0441	[0.0439, 0.0444]	1.02%
$\hat{a}_9$=0.0374	[0.0490, 0.0493]	1.16%
$\hat{a}_{10}$=0.0353	[0.0570, 0.0574]	1.25%
$\hat{a}_{11}$=0.0382	[0.0563, 0.0566]	1.26%
$\hat{a}_{12}$=0.0457	[0.0562, 0.0566]	1.20%
$\hat{a}_{13}$=0.0270	[0.0268, 0.0273]	1.60%
$\hat{a}_{14}$=0.0227	[0.0225, 0.0229]	1.81%
$\hat{a}_{15}$=0.0204	[0.0202, 0.0206]	1.97%
$\hat{a}_{16}$=0.0208	[0.0206, 0.0210]	2.02%
$\hat{a}_{25}$=0.0105	[0.0103, 0.0107]	3.57%
$\hat{a}_{26}$=0.0090	[0.0089, 0.0092]	3.92%
$\hat{a}_{27}$=0.0087	[0.0086, 0.0089]	4.06%
$\hat{a}_{28}$=0.0098	[0.0096, 0.0100]	3.90%
$\hat{a}_{37}$=0.0044	[0.0043, 0.0046]	6.72%
$\hat{a}_{38}$=0.0045	[0.0043, 0.0047]	6.73%
$\hat{a}_{39}$=0.0040	[0.0038, 0.0041]	7.27%
$\hat{a}_{40}$=0.0043	[0.0042, 0.0045]	7.04%
$\hat{a}_{49}$=0.0027	[0.0026, 0.0028]	9.88%
$\hat{a}_{50}$=0.0023	[0.0022, 0.0024]	10.83%
$\hat{a}_{51}$=0.0021	[0.0020, 0.0022]	11.41%
$\hat{a}_{52}$=0.0025	[0.0024, 0.0027]	10.54%
CSA	Fieller's Method	Precision
$\hat{h}$=0.1733	[0.1730, 0.1736]	0.35%
$a\hat{a}_2^A$=0.1242	[0.1240, 0.1244]	0.33%
$\hat{a}_3^A$=0.0956	[0.0954, 0.0957]	0.33%
$\hat{a}_4^A$=0.0768	[0.0766, 0.0769]	0.32%
$\hat{a}_5^A$=0.0627	[0.0626, 0.0628]	0.32%
$\hat{a}_6^A$=0.0530	[0.0529, 0.0530]	0.32%
DD	Delta Method	Precision
$\hat{a}_1^d$=0.2835	[0.2831, 0.2839]	0.28%
$\hat{a}_2^d$=0.1650	[0.1647, 0.1653]	0.40%
$\hat{a}_3^d$=0.1086	[0.1083, 0.1089]	0.51%
$\hat{a}_4^d$=0.0813	[0.0811, 0.0816]	0.59%
$\hat{a}_5^d$=0.0561	[0.0558, 0.0563]	0.72%
$\hat{a}_6^d$=0.0463	[0.0461, 0.0465]	0.80%
$Sum\ of\ S_i$	Delta Method	Precision
$\hat{S}$=5.7699	[5.7599, 5.7798]	0.35%

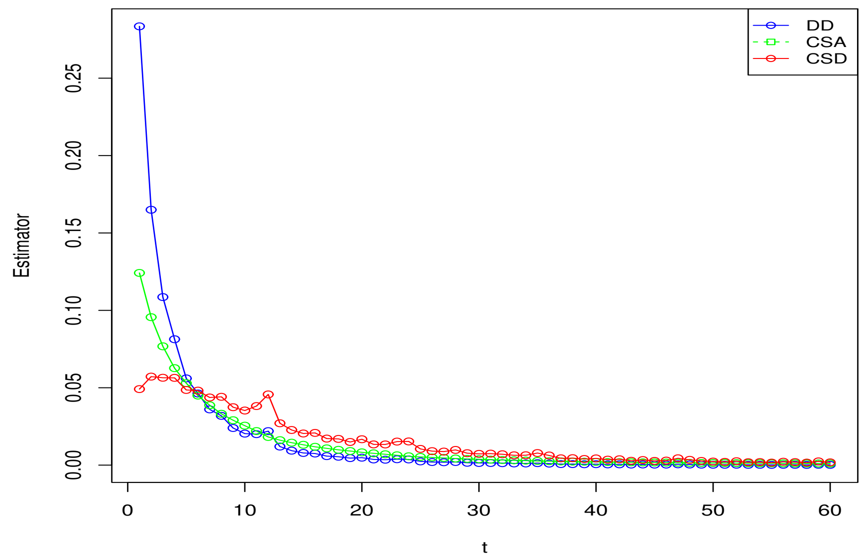

Figure 1: The three distributions estimated from the UK CPI data

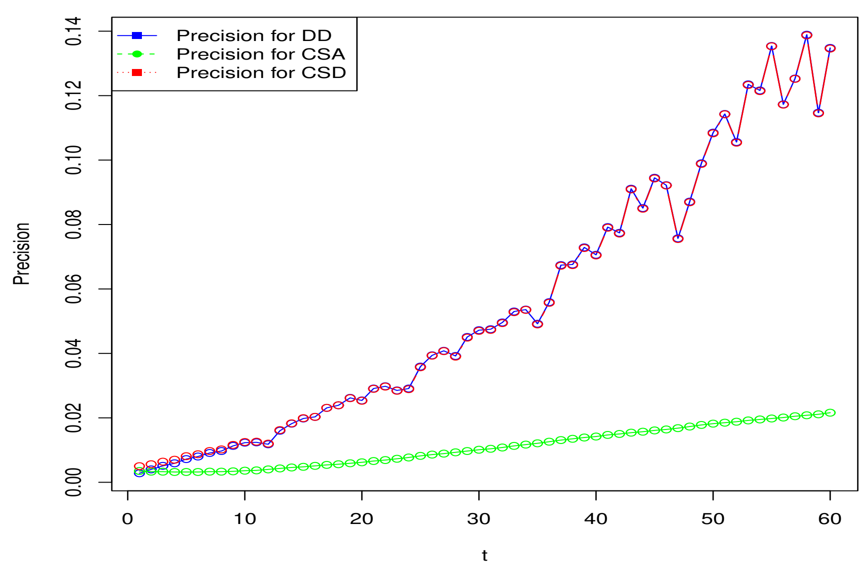

Figure 2: The precision of the three estimators compared.

Reference

Bebu, I., Luta, G., Mathew, T., Kennedy, P. A., and Agan, B. K. (2016). Parametric cost-effectiveness inference with skewed data. *Computational Statistics and Data Analysis*, 94:210–220.

Beyene, J. and Moineddin, R. (2005). Methods for confidence interval estimation of a ratio parameter with application to location quotients. *BMC Medical Research Methodology*, 5(32):1–7.

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Annals of Statistics*, 2:437–453.

Briggs, A. H., Mooney, C. Z., and Wonderling, D. E. (1999). Constructing confidence intervals for cost-effectiveness ratios: an evaluation of parametric and non-parametric techniques using monte carlo simulation. *Statistics in Medicine*, 18:3245–3262.

Buonaccorsi, J. P. (2005). Fieller's theorem. *Encyclopedia of Environmetrics*, pages 1951–1952.

Carvalho, C., Dam, N. A., and Lee, J. W. (2020). The cross-sectional distribution of price stickiness implied by aggregate data. *he Review of Economics and Statistics*, 102(1):162–179.

Cavallo, A. (2016). Scraped data and sticky prices. *Working Paper*.

Cedilnik, A., Košmelj, K., and Blejec, A. (2004). The distribution of the ratio of jointly normal variables. *Metodološki zvezki*, 1(1):99–108.

Cedilnik, A., Košmelj, K., and Blejec, A. (2006). Ratio of two random variables: A note on the existence of its moments. *Metodološki zvezki*, 3(1):1–7.

Cox, C. (1990). Fieller theorem, the likelihood and the delta method. *Biometrics*, 46(3):709–718.

Dixon, H. (2012). A unified framework for using micro-data to compare dynamic time-dependent price-setting models. *BE Journal of Macroeconomics (Contributions)*, 12:1–43.

Dixon, H. and Bihan, H. L. (2012). Generalised taylor and generalised calvo price and wage setting: Micro-evidence with macro implications. *Economic Journal*, 122(560):532–554.

Dixon, H., Luintel, K., and Tian, K. (2020). The impact of the 2008 crisis on UK prices: what we can learn from the CPI microdata. *Oxford Bulletin of Economics*, 82(6):1322–1341.

Dixon, H. and Siciliani, L. (2009). Waiting-time targets in the healthcare sector: how long are we waiting ? *Journal of Health Economics*, 28:1081–1098.

Dixon, H. and Tian, K. (2017). What we can learn about the behavior of firms from the average monthly frequency of price-changes: an application to the UK CPI data. *Oxford Bulletin of Economics*, 79(6):907–932.

Efron, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association*, 76(374):312–319.

Fan, M. and Zhou, X. (2007). A simulation study to compare methods for constructing confidence intervals for the incremental cost-effectiveness ratio. *Statistics in Medicine*, 7:57–77.

Fieller, E. C. (1932). The distribution of the index in a normal bivariate population. *Biometrika*, 24:175–185.

Fieller, E. C. (1940). The biological standardization of insulin. *Supplement to the Journal of the Royal Statistical Society*, 7(1):1–64.

Fieller, E. C. (1954). Some problems in interval estimation. *Journal of Royal Statistical Society. Series B (Methodological)*, 16(2):175–185.

Franz, V. (2007). Ratios: A short guide to confidence limits and proper use. *Working Paper*.

Gardiner, J. C., Huebner, M., Jetton, J., and Bradley, C. J. (2001). On parametric confidence intervals for the cost-effectiveness ratio. *Biometrical Journal*, 43(3):283–296.

Greenwood, M. (1926). The natural duration of cancer. *Reports on Public Health and Medical Subjects*, 33:1–26.

Guiard, V. (1989). Some remarks on the estimation of the ratio of the expectation values of a two-dimensional normal random variable (correction of the theorem of milliken). *Biometika*, 31(6):681–697.

Kaplan, E. L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association*, 53:457–481.

Marsaglia, G. (1965). Ratios of normal variables and ratios of sums of uniform variables. *Journal of the American Statistical Association*, 60(309):193–204.

Nakamura, E. and Steinsson, J. (2008). Five facts about prices: A re-evaluation of menu cost models. *Quarterly Journal of Economics*, 123(4):1415–1464.

Polsky, D., Glick, H. A., Willke, R., and Schulman, K. (1997). Confidence intervals for cost-effectiveness ratios a comparison of four methods. *Health Economics*, 6(3):243–252.

Tian, M. and Dixon, H. (2022). The variances of non-parametric estimates of the cross-sectional distribution of durations. *Econometric Reviews*, 41(10):1243–1264.

Tsai, W. Y., Jewell, N. P., and Wang, M. (1987). A note on the product-limit estimator under right censoring and left truncation. *Biometrika*, 74(4):883–886.

Wang, H. and Zhao, H. (2008). A study on confidence intervals for incremental cost-effectiveness ratios. *Biometrical Journal*, 50:505–514.

The Confidence Interval of the Cross-Sectional Distribution of Durations: Online Appendix

Huw Dixon and Maoshan Tian

In this online appendix we provide the full range of estimates and confidence intervals up to and including 60 months of the estimates and the CIs of CSD, DD and CSA. For CSD and CSA we show the CIs calculated using the Fieller and delta method. The results for CSD are in Table OA1, CSA in Table OA2 and DD in Table OA3. We also depict the hazard function estimates in Figure 1 for all durations, and the estimates and CIs in figures 1-3 for the 5th year (months 49-60) for each of the three distributions. Table OA1 shows the CSD estimators of UK CPI price-quotes from 1998m12 to 2017m1. Since Fieller's method and delta method provide nearly the same CIs when there are only 4 decimal places included, so the column CI in table OA1 is the results from both Fieller's method and delta method. The column P in table OA1 is the precision ratio from Fieller's method and delta method. Tale OA2 shows the CSA estimators of UK CPI price-quotes from 1998m12 to 2017m1. We also combine results of Fieller's method and delta method for the CI and P since they are nearly the same if we only report 4 decimal places (the differences exists when we include 7 or 8 decimal places.) Table OA3 shows the DD estimators of UK CPI price-quotes from 1998m12 to 2017m1. We only use delta method to calculate the CI and P for DD estimators.

Table OA1 The 90% CI of CSD for CPI Data From 1998m12 to 2017m1

CSD	CI	P
$\hat{a}_1$= 0.0491	[0.0490, 0.0493]	0.0050
$\hat{a}_2$= 0.0572	[0.0570, 0.0574]	0.0056
$\hat{a}_3$= 0.0565	[0.0563, 0.0566]	0.0064
$\hat{a}_4$= 0.0564	[0.0562, 0.0566]	0.0070
$\hat{a}_5$= 0.0486	[0.0484, 0.0488]	0.0081
$\hat{a}_6$= 0.0481	[0.0479, 0.0483]	0.0087
$\hat{a}_7$= 0.0437	[0.0435, 0.0439]	0.0097
$\hat{a}_8$= 0.0441	[0.0439, 0.0444]	0.0102
$\hat{a}_9$= 0.0374	[0.0372, 0.0376]	0.0116
$\hat{a}_{10}$= 0.0353	[0.0351, 0.0355]	0.0125
$\hat{a}_{11}$= 0.0382	[0.0379, 0.0384]	0.0126
$\hat{a}_{12}$= 0.0457	[0.0454, 0.0460]	0.0120
$\hat{a}_{13}$= 0.0270	[0.0268, 0.0273]	0.0162
$\hat{a}_{14}$= 0.0227	[0.0225, 0.0229]	0.0183
$\hat{a}_{15}$= 0.0204	[0.0203, 0.0206]	0.0199
$\hat{a}_{16}$= 0.0208	[0.0206, 0.0210]	0.0203
$\hat{a}_{17}$= 0.0171	[0.0170, 0.0173]	0.0231
$\hat{a}_{18}$= 0.0170	[0.0168, 0.0172]	0.0239
$\hat{a}_{19}$= 0.0150	[0.0148, 0.0151]	0.0261
$\hat{a}_{20}$= 0.0167	[0.0165, 0.0169]	0.0253
$\hat{a}_{21}$= 0.0134	[0.0132, 0.0136]	0.0290
$\hat{a}_{22}$= 0.0134	[0.0132, 0.0136]	0.0297
$\hat{a}_{23}$= 0.0152	[0.0150, 0.0154]	0.0284
$\hat{a}_{24}$= 0.0153	[0.0151, 0.0155]	0.0289
$\hat{a}_{25}$= 0.0105	[0.0103, 0.0107]	0.0357
$\hat{a}_{26}$= 0.0090	[0.0089, 0.0092]	0.0393
$\hat{a}_{27}$= 0.0087	[0.0086, 0.0089]	0.0407
$\hat{a}_{28}$= 0.0098	[0.0096, 0.0100]	0.0390
$\hat{a}_{29}$= 0.0077	[0.0075, 0.0079]	0.0449
$\hat{a}_{30}$= 0.0073	[0.0071, 0.0075]	0.0470
$\hat{a}_{31}$= 0.0074	[0.0073, 0.0076]	0.0473
$\hat{a}_{32}$= 0.0070	[0.0069, 0.0072]	0.0494
$\hat{a}_{33}$= 0.0064	[0.0062, 0.0065]	0.0528
$\hat{a}_{34}$= 0.0064	[0.0062, 0.0066]	0.0535
$\hat{a}_{35}$= 0.0078	[0.0076, 0.0080]	0.0490
$\hat{a}_{36}$= 0.0062	[0.0061, 0.0064]	0.0557
$\hat{a}_{37}$= 0.0044	[0.0043, 0.0046]	0.0671
$\hat{a}_{38}$= 0.0045	[0.0043, 0.0047]	0.0674
$\hat{a}_{39}$= 0.0040	[0.0038, 0.0041]	0.0727
$\hat{a}_{40}$= 0.0043	[0.0042, 0.0045]	0.0704
$\hat{a}_{41}$= 0.0035	[0.0034, 0.0037]	0.0790
$\hat{a}_{42}$= 0.0038	[0.0036, 0.0039]	0.0772
$\hat{a}_{43}$= 0.0028	[0.0027, 0.0029]	0.0909
$\hat{a}_{44}$= 0.0033	[0.0031, 0.0034]	0.0849
$\hat{a}_{45}$= 0.0027	[0.0026, 0.0029]	0.0943
$\hat{a}_{46}$= 0.0029	[0.0028, 0.0031]	0.0921
$\hat{a}_{47}$= 0.0044	[0.0043, 0.0046]	0.0755
$\hat{a}_{48}$= 0.0034	[0.0033, 0.0036]	0.0869
$\hat{a}_{49}$= 0.0027	[0.0026, 0.0028]	0.0988
$\hat{a}_{50}$= 0.0023	[0.0022, 0.0024]	0.1083
$\hat{a}_{51}$= 0.0021	[0.0020, 0.0022]	0.1142
$\hat{a}_{52}$= 0.0025	[0.0024, 0.0027]	0.1054
$\hat{a}_{53}$= 0.0019	[0.0018, 0.0020]	0.1233
$\hat{a}_{54}$= 0.0020	[0.0019, 0.0021]	0.1214
$\hat{a}_{55}$= 0.0016	[0.0015, 0.0017]	0.1353
$\hat{a}_{56}$= 0.0022	[0.0021, 0.0023]	0.1172
$\hat{a}_{57}$= 0.0020	[0.0018, 0.0021]	0.1252
$\hat{a}_{58}$= 0.0016	[0.0015, 0.0017]	0.1387
$\hat{a}_{59}$= 0.0024	[0.0023, 0.0026]	0.1145
$\hat{a}_{60}$= 0.0018	[0.0017, 0.0019]	0.1346
$\hat{a}_{61}$= 0.0621	[0.0614, 0.0627]	0.0219

Table OA2The 90% CI of CSA for CPI Data From 1998m12 to 2017m1

CSA	CI	P
$\hat{a}_1^A$= 0.1733	[0.1730, 0.1736]	0.0035
$\hat{a}_2^A$= 0.1242	[0.1240, 0.1244]	0.0033
$\hat{a}_3^A$= 0.0956	[0.0954, 0.0957]	0.0033
$\hat{a}_4^A$= 0.0768	[0.0766, 0.0769]	0.0032
$\hat{a}_5^A$= 0.0627	[0.0626, 0.0628]	0.0032
$\hat{a}_6^A$= 0.0530	[0.0529, 0.0530]	0.0032
$\hat{a}_7^A$= 0.0449	[0.0449, 0.0450]	0.0033
$\hat{a}_8^A$= 0.0387	[0.0386, 0.0388]	0.0033
$\hat{a}_9^A$= 0.0332	[0.0331, 0.0332]	0.0034
$\hat{a}_{10}^A$= 0.0290	[0.0290, 0.0291]	0.0036
$\hat{a}_{11}^A$= 0.0255	[0.0254, 0.0255]	0.0037
$\hat{a}_{12}^A$= 0.0220	[0.0220, 0.0221]	0.0040
$\hat{a}_{13}^A$= 0.0182	[0.0182, 0.0182]	0.0043
$\hat{a}_{14}^A$= 0.0161	[0.0161, 0.0162]	0.0046
$\hat{a}_{15}^A$= 0.0145	[0.0145, 0.0145]	0.0048
$\hat{a}_{16}^A$= 0.0131	[0.0131, 0.0132]	0.0051
$\hat{a}_{17}^A$= 0.0118	[0.0118, 0.0119]	0.0054
$\hat{a}_{18}^A$= 0.0108	[0.0108, 0.0109]	0.0056
$\hat{a}_{19}^A$= 0.0099	[0.0099, 0.0099]	0.0059
$\hat{a}_{20}^A$= 0.0091	[0.0091, 0.0091]	0.0062
$\hat{a}_{21}^A$= 0.0083	[0.0082, 0.0083]	0.0066
$\hat{a}_{22}^A$= 0.0076	[0.0076, 0.0077]	0.0069
$\hat{a}_{23}^A$= 0.0070	[0.0070, 0.0070]	0.0073
$\hat{a}_{24}^A$= 0.0064	[0.0063, 0.0064]	0.0077
$\hat{a}_{25}^A$= 0.0057	[0.0057, 0.0057]	0.0082
$\hat{a}_{26}^A$= 0.0053	[0.0053, 0.0053]	0.0085
$\hat{a}_{27}^A$= 0.0050	[0.0049, 0.0050]	0.0089
$\hat{a}_{28}^A$= 0.0046	[0.0046, 0.0047]	0.0093
$\hat{a}_{29}^A$= 0.0043	[0.0043, 0.0043]	0.0097
$\hat{a}_{30}^A$= 0.0040	[0.0040, 0.0040]	0.0101
$\hat{a}_{31}^A$= 0.0038	[0.0038, 0.0038]	0.0104
$\hat{a}_{32}^A$= 0.0035	[0.0035, 0.0036]	0.0108
$\hat{a}_{33}^A$= 0.0033	[0.0033, 0.0033]	0.0113
$\hat{a}_{34}^A$= 0.0031	[0.0031, 0.0031]	0.0117
$\hat{a}_{35}^A$= 0.0029	[0.0029, 0.0029]	0.0121
$\hat{a}_{36}^A$= 0.0027	[0.0027, 0.0027]	0.0126
$\hat{a}_{37}^A$= 0.0025	[0.0025, 0.0026]	0.0131
$\hat{a}_{38}^A$= 0.0024	[0.0024, 0.0024]	0.0135
$\hat{a}_{39}^A$= 0.0023	[0.0023, 0.0023]	0.0139
$\hat{a}_{40}^A$= 0.0022	[0.0022, 0.0022]	0.0142
$\hat{a}_{41}^A$= 0.0021	[0.0021, 0.0021]	0.0147
$\hat{a}_{42}^A$= 0.0020	[0.0020, 0.0020]	0.0150
$\hat{a}_{43}^A$= 0.0019	[0.0019, 0.0019]	0.0154
$\hat{a}_{44}^A$= 0.0018	[0.0018, 0.0019]	0.0157
$\hat{a}_{45}^A$= 0.0018	[0.0018, 0.0018]	0.0161
$\hat{a}_{46}^A$= 0.0017	[0.0017, 0.0017]	0.0164
$\hat{a}_{47}^A$= 0.0016	[0.0016, 0.0017]	0.0168
$\hat{a}_{48}^A$= 0.0016	[0.0015, 0.0016]	0.0173
$\hat{a}_{49}^A$= 0.0015	[0.0015, 0.0015]	0.0178
$\hat{a}_{50}^A$= 0.0014	[0.0014, 0.0014]	0.0182
$\hat{a}_{51}^A$= 0.0014	[0.0014, 0.0014]	0.0185
$\hat{a}_{52}^A$= 0.0013	[0.0013, 0.0014]	0.0188
$\hat{a}_{53}^A$= 0.0013	[0.0013, 0.0013]	0.0192
$\hat{a}_{54}^A$= 0.0013	[0.0012, 0.0013]	0.0195
$\hat{a}_{55}^A$= 0.0012	[0.0012, 0.0012]	0.0198
$\hat{a}_{56}^A$= 0.0012	[0.0012, 0.0012]	0.0201
$\hat{a}_{57}^A$= 0.0012	[0.0011, 0.0012]	0.0205
$\hat{a}_{58}^A$= 0.0011	[0.0011, 0.0011]	0.0208
$\hat{a}_{59}^A$= 0.0011	[0.0011, 0.0011]	0.0211
$\hat{a}_{60}^A$= 0.0010	[0.0010, 0.0011]	0.0216
$\hat{a}_{61}^A$= 0.0010	[0.0010, 0.0010]	0.0219

Table OA3The 90% CI of DD for CPI Data From 1998m12 to 2017m1

DD	CI	P
$\hat{a}_1^d = 0.2835$	[0.2831, 0.2839]	0.0028
$\hat{a}_2^d = 0.1650$	[0.1647, 0.1653]	0.0040
$\hat{a}_3^d = 0.1086$	[0.1083, 0.1089]	0.0051
$\hat{a}_4^d = 0.0813$	[0.0811, 0.0816]	0.0059
$\hat{a}_5^d = 0.0561$	[0.0558, 0.0563]	0.0072
$\hat{a}_6^d = 0.0463$	[0.0461, 0.0465]	0.0080
$\hat{a}_7^d = 0.0360$	[0.0359, 0.0362]	0.0091
$\hat{a}_8^d = 0.0318$	[0.0317, 0.0320]	0.0097
$\hat{a}_9^d = 0.0240$	[0.0239, 0.0241]	0.0113
$\hat{a}_{10}^d = 0.0204$	[0.0202, 0.0205]	0.0123
$\hat{a}_{11}^d = 0.0200$	[0.0199, 0.0201]	0.0124
$\hat{a}_{12}^d = 0.0220$	[0.0218, 0.0221]	0.0118
$\hat{a}_{13}^d = 0.0120$	[0.0119, 0.0121]	0.0160
$\hat{a}_{14}^d = 0.0094$	[0.0093, 0.0094]	0.0182
$\hat{a}_{15}^d = 0.0079$	[0.0078, 0.0079]	0.0198
$\hat{a}_{16}^d = 0.0075$	[0.0074, 0.0076]	0.0203
$\hat{a}_{17}^d = 0.0058$	[0.0058, 0.0059]	0.0231
$\hat{a}_{18}^d = 0.0054$	[0.0054, 0.0055]	0.0239
$\hat{a}_{19}^d = 0.0045$	[0.0045, 0.0046]	0.0262
$\hat{a}_{20}^d = 0.0048$	[0.0048, 0.0049]	0.0254
$\hat{a}_{21}^d = 0.0037$	[0.0036, 0.0037]	0.0291
$\hat{a}_{22}^d = 0.0035$	[0.0035, 0.0036]	0.0298
$\hat{a}_{23}^d = 0.0038$	[0.0038, 0.0039]	0.0285
$\hat{a}_{24}^d = 0.0037$	[0.0036, 0.0037]	0.0291
$\hat{a}_{25}^d = 0.0024$	[0.0024, 0.0025]	0.0359
$\hat{a}_{26}^d = 0.0020$	[0.0020, 0.0020]	0.0394
$\hat{a}_{27}^d = 0.0019$	[0.0018, 0.0019]	0.0408
$\hat{a}_{28}^d = 0.0020$	[0.0020, 0.0021]	0.0392
$\hat{a}_{29}^d = 0.0015$	[0.0015, 0.0016]	0.0451
$\hat{a}_{30}^d = 0.0014$	[0.0014, 0.0014]	0.0472
$\hat{a}_{31}^d = 0.0014$	[0.0013, 0.0014]	0.0475
$\hat{a}_{32}^d = 0.0013$	[0.0012, 0.0013]	0.0496
$\hat{a}_{33}^d = 0.0011$	[0.0011, 0.0011]	0.0530
$\hat{a}_{34}^d = 0.0011$	[0.0011, 0.0011]	0.0536
$\hat{a}_{35}^d = 0.0013$	[0.0013, 0.0013]	0.0492
$\hat{a}_{36}^d = 0.0010$	[0.0010, 0.0010]	0.0559
$\hat{a}_{37}^d = 0.0007$	[0.0007, 0.0007]	0.0674
$\hat{a}_{38}^d = 0.0007$	[0.0007, 0.0007]	0.0676
$\hat{a}_{39}^d = 0.0006$	[0.0006, 0.0006]	0.0729
$\hat{a}_{40}^d = 0.0006$	[0.0006, 0.0006]	0.0706
$\hat{a}_{41}^d = 0.0005$	[0.0005, 0.0005]	0.0793
$\hat{a}_{42}^d = 0.0005$	[0.0005, 0.0005]	0.0774
$\hat{a}_{43}^d = 0.0004$	[0.0004, 0.0004]	0.0911
$\hat{a}_{44}^d = 0.0004$	[0.0004, 0.0004]	0.0851
$\hat{a}_{45}^d = 0.0003$	[0.0003, 0.0004]	0.0945
$\hat{a}_{46}^d = 0.0004$	[0.0003, 0.0004]	0.0922
$\hat{a}_{47}^d = 0.0005$	[0.0005, 0.0006]	0.0757
$\hat{a}_{48}^d = 0.0004$	[0.0004, 0.0004]	0.0871
$\hat{a}_{49}^d = 0.0003$	[0.0003, 0.0003]	0.0990
$\hat{a}_{50}^d = 0.0003$	[0.0003, 0.0003]	0.1085
$\hat{a}_{51}^d = 0.0002$	[0.0002, 0.0003]	0.1143
$\hat{a}_{52}^d = 0.0003$	[0.0003, 0.0003]	0.1056
$\hat{a}_{53}^d = 0.0002$	[0.0002, 0.0002]	0.1234
$\hat{a}_{54}^d = 0.0002$	[0.0002, 0.0002]	0.1216
$\hat{a}_{55}^d = 0.0002$	[0.0002, 0.0002]	0.1354
$\hat{a}_{56}^d = 0.0002$	[0.0002, 0.0002]	0.1173
$\hat{a}_{57}^d = 0.0002$	[0.0002, 0.0002]	0.1253
$\hat{a}_{58}^d = 0.0002$	[0.0002, 0.0002]	0.1388
$\hat{a}_{59}^d = 0.0002$	[0.0002, 0.0003]	0.1147
$\hat{a}_{60}^d = 0.0002$	[0.0002, 0.0002]	0.1348
$\hat{a}_{61}^d = 0.0059$	[0.0058, 0.0059]	0.0230

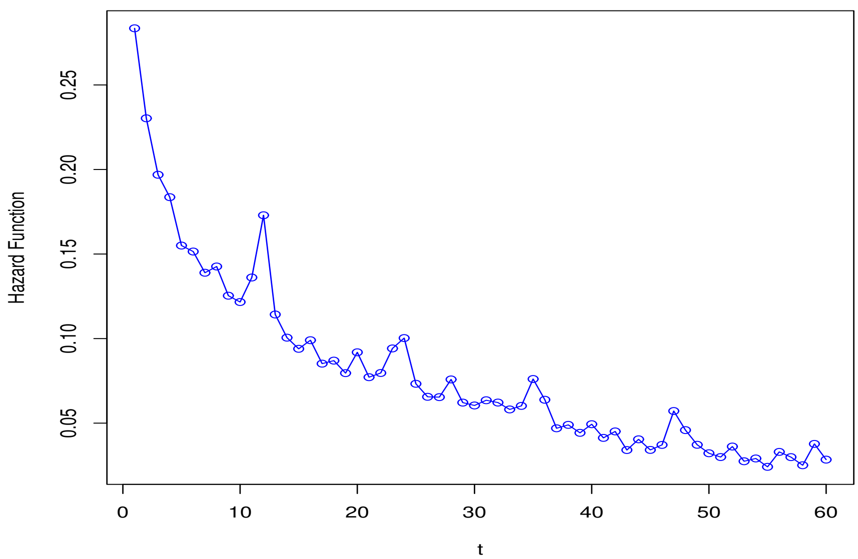

Figure 1: The Hazard Function of UK Micro-CPI Data

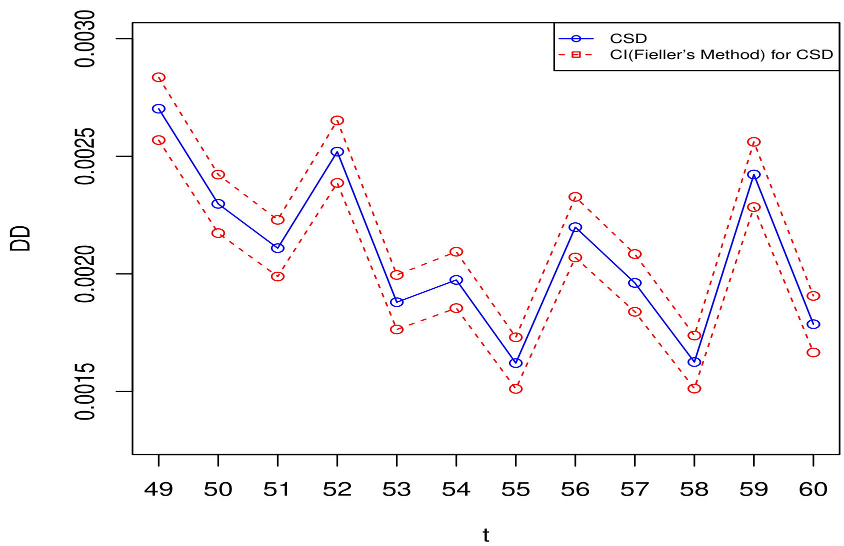

Figure 2: The CSD Estimators and Its CIs of UK Micro-CPI Data for the Fifth Year

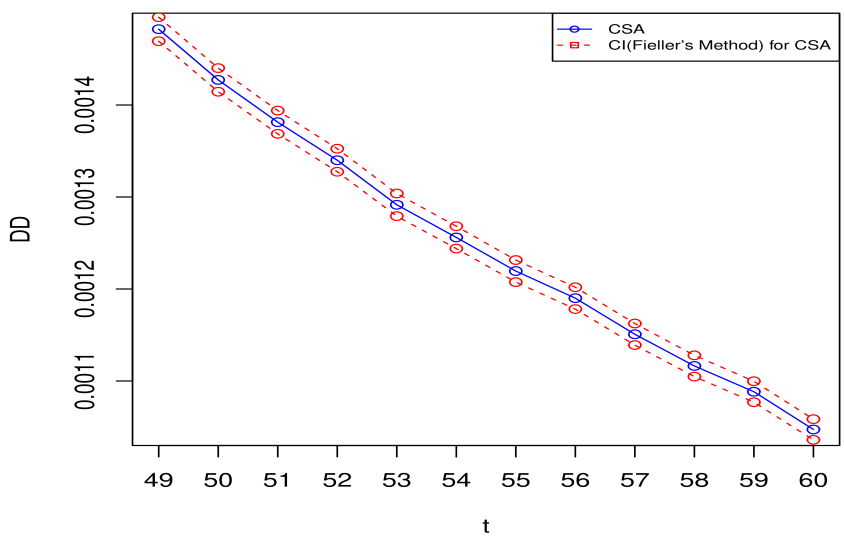

Figure 3: The CSA Estimators and Its CIs of UK Micro-CPI Data for the Fifth Year

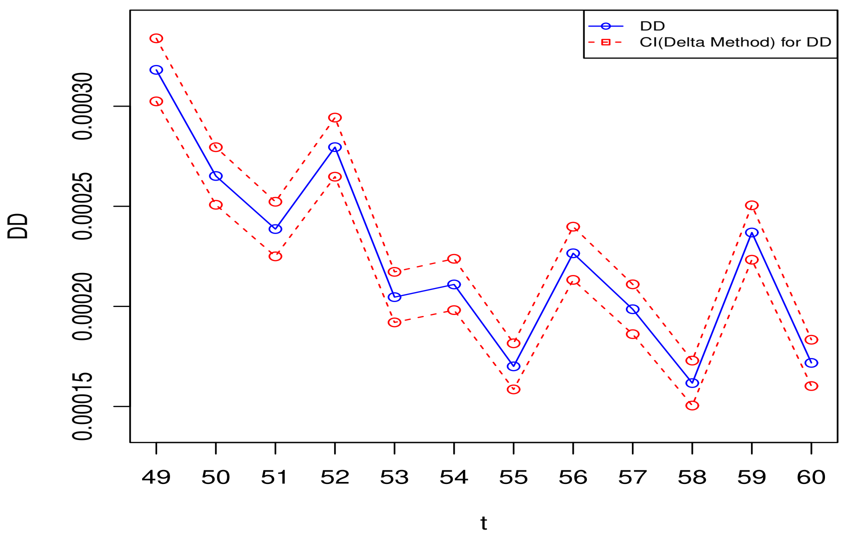

Figure 4: The DD Estimators and Its CIs of UK Micro-CPI Data for the Fifth Year